Theta functions and weighted theta functions of Euclidean lattices, with some applications

Noam D. Elkies[1]

March, 2009

0. Introduction and overview

[. . .]

1. Lattices in R^n: basic terminology, notations, and examples

By "Euclidean space" of dimension n we mean a real vector space of dimension n, equipped with a positive-definite inner product $\langle \cdot, \cdot \rangle$. We usually call such a space "$\mathbf{R}^n$" even when there is no distinguished choice of coordinates.

A *lattice* in $\mathbf{R}^n$ is a discrete co-compact subgroup $L \subset \mathbf{R}^n$, that is, a discrete subgroup such that the quotient $\mathbf{R}^n/L$ is compact (and thus necessarily homeomorphic with the n-torus $(\mathbf{R}/\mathbf{Z})^n$). As an abstract group L is thus isomorphic with the free abelian group $\mathbf{Z}^n$ of rank n. Therefore L is determined by the images, call them $v_1, \ldots, v_n$, of the standard generators of $\mathbf{Z}^n$ under a group isomorphism $\mathbf{Z}^n \xrightarrow{\sim} L$. We say the v_i *generate*, or are *generators* of, L: each vector in L can be written as $\sum_{i=1}^n a_i v_i$ for some *unique* integers $a_1, \ldots, a_n$. Vectors $v_1, \ldots, v_n \in \mathbf{R}^n$ generate a lattice if and only if they constitute an $\mathbf{R}$-linear basis for $\mathbf{R}^n$, and then L is the $\mathbf{Z}$-span of this basis. For instance, the $\mathbf{Z}$-span of the standard orthonormal basis $e_1, \ldots, e_n$ of $\mathbf{R}^n$ is the lattice $\mathbf{Z}^n$. This more concrete definition is better suited for explicit computation, but less canonical because most lattices have no canonical choice of generators even up to isometries of $\mathbf{R}^n$.

An equivalent approach defines L as a free abelian group of rank n with a positive-definite symmetric pairing; the Euclidean space is then recovered as $L \otimes_{\mathbf{Z}} \mathbf{R}$, which inherits a symmetric bilinear pairing from L. With this approach we must be careful about the definition of a "positive-definite pairing" so that the extension to $L \otimes_{\mathbf{Z}} \mathbf{R}$ remains positive-definite. For most of the lattices we consider, the pairing takes values in $\mathbf{Q}$, and then the usual definition for inner products suffices:

$$v \in L,\ v \neq 0 \implies \langle v, v \rangle > 0. \tag{1}$$

[1]Department of Mathematics, Harvard University, Cambridge, MA 02138 USA; e-mail: (elkies@math.harvard.edu). Supported in part by NSF grant DMS-0501029.

But for a general $\mathbf{R}$-valued pairing, (1) guarantees only that the pairing on $L \otimes_{\mathbf{Z}} \mathbf{R}$ is positive semidefinite, not necessarily positive definite: a standard counterexample is $L = \mathbf{Z}^2$ with

$$\langle v, w \rangle = (v_1 - tv_2)(w_1 - tw_2)$$

for some irrational t, when the nonzero vector $x = (t, 1) \in L \otimes_{\mathbf{Z}} \mathbf{R}$ satisfies $\langle x, x \rangle = 0$. For this general case we say the pairing $\langle \cdot, \cdot \rangle$ is "positive-definite" if it has a positive-definite Gram matrix A. The *Gram matrix* of $\langle \cdot, \cdot \rangle$ with respect to generators $v_1, \ldots, v_n$ of L is the symmetric matrix with (i, j) entry $A_{ij} = \langle v_i, v_j \rangle$. The Gram matrix depends on the choice of generators, but if $w_1, \ldots, w_n$ are any other generators then each $w_k = \sum_{i=1}^n b_{ik} v_i$ for some integers b_{ik} forming a matrix B of determinant ± 1, and the Gram matrix with respect to $w_1, \ldots, w_n$ is $B^{\mathsf{T}} A B$, which is positive-definite if and only if A is. If $v = m_1 v_1 + \cdots + m_n v_n$ for some $m = (m_1, \ldots, m_n) \in \mathbf{Z}^n$ then $\langle v, v \rangle$ is the value at m of the *quadratic form*

$$(m, Am^{\mathsf{T}}) = \sum_{i=1}^n \sum_{j=1}^n A_{ij} m_i m_j = \sum_{i=1}^n \sum_{j=1}^n \langle v_i, v_j \rangle m_i m_j, \tag{2}$$

a homogeneous polynomial of degree 2 in n variables.

We next recall some further invariants[2] of L and the corresponding properties of the Gram matrix. The *discriminant* of L is

$$\operatorname{disc} L = (\operatorname{Vol}(\mathbf{R}^n / L))^2 = \det A; \tag{3}$$

in particular, $\det A$ is independent of the choice of generators, which we can also verify directly: if $\det B = \pm 1$ then $\det A = \det B^{\mathsf{T}} A B$. The volume $\sqrt{\operatorname{disc} L}$ of the torus $\mathbf{R}^n / L$ is known as the *covolume* of L. A lattice of discriminant 1 is said to be *unimodular*. The *dual lattice* L^* is defined by

$$L^* = \{v^* \in \mathbf{R}^n \mid \forall v \in L, \langle v, v' \rangle \in \mathbf{Z}\}. \tag{4}$$

If L is the $\mathbf{Z}$-span of $v_1, \ldots, v_n$ with Gram matrix A, then L^* is the $\mathbf{Z}$-span of the dual basis $v_1^*, \ldots, v_n^*$ with Gram matrix $(A^{\mathsf{T}})^{-1}$; in particular

$$\operatorname{disc} L^* = (\operatorname{disc} L)^{-1}. \tag{5}$$

We say L is *integral* if $\langle v, v' \rangle \in \mathbf{Z}$ for all $v, v' \in L$; equivalently, if $L \subseteq L^*$. In this case L^*/L is a finite group with $\#(L^*/L) = \operatorname{disc}(L)$. In particular, $L = L^*$ if and only if L is integral and unimodular; we naturally say such a lattice is *self-dual*.

[2] These are "invariant" in the sense that they do not depend on a choice of generators, nor on other extrinsic features such as an embedding in $\mathbf{R}^n$. The Gram matrix depends on the choice of generators, and is thus not an invariant: different Gram matrices may give rise to the same lattice.

The basic example of a lattice is $\mathbf{Z}$ with the pairing $\langle x, y\rangle = xy$; this lattice is self-dual, and is the unique unimodular lattice in the 1-dimensional Euclidean space $\mathbf{R}$. More generally, for every real $D > 0$ there is a unique lattice of discriminant D in $\mathbf{R}$, namely $D^{1/2}\mathbf{Z}$, or equivalently $\mathbf{Z}$ with the pairing $\langle x, y\rangle = Dxy$ instead; this lattice is integral if and only if $D \in \mathbf{Z}$.

We next give some constructions of new lattices from old. A subgroup L' of finite index in a lattice L in $\mathbf{R}^n$ is itself a lattice in $\mathbf{R}^n$ whose dual contains L^*. Comparing covolumes, we see that $[L'^* : L^*] = [L : L']$; in fact much more is true: the inner product on $\mathbf{R}^n$ induces a perfect pairing $(L'^*/L^*) \times (L/L') \to \mathbf{Q}/\mathbf{Z}$ on the quotient subgroups, so in particular these subgroups are isomorphic, albeit not in general canonically isomorphic.

If $v_1, \ldots, v_n$ generate L and $w_1, \ldots, w_n$ generate L', then $[L : L'] = |\det B|$, where B is the integer matrix formed by the coefficients b_{ik} of the expansions $w_k = \sum_{i=1}^n b_{ik} v_i$. If A is the Gram matrix of L with respect to the v_i, then as before $\det B^{\mathsf{T}} A B$ is the Gram matrix of L' with respect to the w_k. Therefore

$$\operatorname{disc} L' = [L : L']^2 \operatorname{disc} L, \tag{6}$$

an identity that can also be obtained from the first equality in (3) because

$$\operatorname{Vol}(\mathbf{R}^n/L') = [L : L'] \operatorname{Vol}(\mathbf{R}^n/L).$$

If L is integral then so is L'. In the other direction, if L' is integral, let G be any set of generators of L/L', and $\tilde{G}$ an arbitrary lift of G to a subset of L; then L is integral if and only if $L \subset L'^*$ and $\langle \tilde{g}, \tilde{g}' \rangle \in \mathbf{Z}$ for all $\tilde{g}, \tilde{g}' \in \tilde{G}$.

If $n = n_1 + n_2$ and L_1 and L_2 are lattices in $\mathbf{R}^{n_1}$ and $\mathbf{R}^{n_2}$ respectively, then the direct sum

$$L = L_1 \oplus L_2 := \{(v_1, v_2) \mid v_1 \in L_1, v_2 \in L_2\} \tag{7}$$

is a lattice in $\mathbf{R}^n$ of discriminant $\operatorname{disc}(L_1) \operatorname{disc}(L_2)$; if A_i ($i = 1, 2$) is a Gram matrix for L_i then L has a block-diagonal Gram matrix with blocks A_1, A_2. The dual of $L_1 \oplus L_2$ is $L_1^* \oplus L_2^*$. The direct sum L is integral if and only if L_1 and L_2 are integral, and self-dual if and only if L_1 and L_2 are self-dual. For example, $\mathbf{Z}^2 = \mathbf{Z} \oplus \mathbf{Z} \subset \mathbf{R}^2$ is self-dual; iterating the construction yields the self-dual lattice $\mathbf{Z}^n \subset \mathbf{R}^n$ for each $n = 1, 2, 3, \ldots$.

Of course once $n > 1$ this self-dual lattice is no longer unique, because we can obtain uncountably many others by applying an orthogonal linear transformation to $\mathbf{R}^n$; but the resulting lattices are isomorphic. It is known that for $n \leq 7$ every self-dual lattice is isomorphic with $\mathbf{Z}^n$ (see Proposition 7 below for one approach to this result), and for every n there are only finitely many isomorphism classes.

But for large n the number of isomorphism classes grows rapidly, exceeding $(cn)^{n^2}$ for some positive c. We conclude this section by constructing the self-dual lattice $E_8 \subset \mathbf{R}^8$ and showing that $E_8 \ncong \mathbf{Z}^8$.

For $n = 1, 2, 3, \ldots$, define a lattice $D_n \in \mathbf{R}^n$ by

$$D_n = \big\{(x_1, \ldots, x_n) \in \mathbf{Z}^n : \sum_{i=1}^{n} x_i \equiv 0 \bmod 2\big\}, \tag{8}$$

a sublattice[3] of $\mathbf{Z}^n$ of index 2. An explicit $\mathbf{Z}$-basis consists of $e_i + e_{i+1}$ for $0 < i < n$ together with $2e_1$. The dual lattice D_n^* is the union of $\mathbf{Z}^n$ and the translate of $\mathbf{Z}^n$ by the half-lattice vector

$$h := (1, 1, \ldots, 1)/2 = \frac{1}{2}\sum_{i=1}^{n} e_i. \tag{9}$$

The 4-element quotient group D_n^*/D_n is cyclic if n is odd, and of exponent 2 when n is even. In the latter case, $2h \in D_n$, so

$$D_n^+ := D_n \cup (D_n + h) \tag{10}$$

is a lattice, which is unimodular because $[D_n^+ : D_n] = 2 = [\mathbf{Z}^n : D_n]$. An explicit basis consists of $e_i + e_{i+1}$ for $0 < i < n-1$ together with $2e_1$ and h. This lattice is integral, and thus self-dual, if and only if $4|n$, because $\langle h, h\rangle = n/4$. Having claimed that for $n < 8$ every unimodular lattice in $\mathbf{R}^n$ is isomorphic with $\mathbf{Z}^n$, we should have $D_4^+ \cong \mathbf{Z}^4$, and indeed h, $h - e_1$, $h - e_2$, $h - e_3$ are orthonormal generators for D_4^+. But if $n \neq 4$ then $D_n^+ \ncong \mathbf{Z}^n$, because D_n^+ contains no vectors x with $\langle x, x\rangle = 1$. In particular, D_8^+ is a self-dual lattice not isomorphic with $\mathbf{Z}^8$. The lattice D_8^+ is usually called E_8; it has some remarkable properties (some of which we shall see later) not shared by D_n^+ for $n \neq 8$. It is known (see for instance [CS2, p.49 and Table 16.7]) that for each $n \leq 13$ every self-dual lattice in $\mathbf{R}^n$ is isomorphic with one of $\mathbf{Z}^n$, $\mathbf{Z}^{n-8} \oplus E_8$ (for $n \geq 8$), or $\mathbf{Z}^{n-12} \oplus D_{12}^+$ (for $n = 12$ or $n = 13$). For $n = 8$ this means every self-dual lattice in $\mathbf{R}^8$ is isomorphic with either $\mathbf{Z}^8$ or E_8; see Exercise 3.5 below (and Exercise 3.6 for the classification for all $n \leq 15$).

[3]Warning: since we regard any lattice L as living in a Euclidean space $\mathbf{R}^n = L \otimes \mathbf{R}$, any subgroup L' is also a lattice in some Euclidean space, but possibly a proper subspace of $\mathbf{R}^n$ — indeed L' is a lattice in $\mathbf{R}^n$ if and only if the index is $[L : L']$ is finite. But we call L' a "sublattice" whether or not it is of finite index; thus a "sublattice of a lattice in $\mathbf{R}^n$" is not necessarily a "lattice in $\mathbf{R}^n$" in our sense.

Exercises

1.1 i) Every lattice L in $\mathbf{R}^n$ has, for each integer $m > 0$, a sublattice mL. How are the Gram matrix, discriminant, and dual lattice of mL related to those of L?
ii) For any lattice L in $\mathbf{R}^n$ and any positive $\alpha \in \mathbf{R}$, define $L\langle\alpha\rangle$ to be the lattice obtained from L by multiplying the pairing by α. For example, $mL \cong L\langle m^2\rangle$. How are the Gram matrix, discriminant, and dual lattice of $L\langle\alpha\rangle$ related to those of L?

1.2 i) Verify that the 4-element group D_n^*/D_n is cyclic if and only if n is odd, and that D_n is irreducible (not the direct sum of two sublattices of positive dimension) except for $D_2 \cong \mathbf{Z}\langle 2\rangle \oplus \mathbf{Z}\langle 2\rangle (\cong \mathbf{Z}^2\langle 2\rangle)$. [For the irreducibility of D_n for $n \geq 3$ you might consider the vectors $x \in D_n$ with $\langle x, x\rangle = 2$, which is the smallest value of $\langle x, x\rangle$ for nonzero $x \in D_n$.]
ii) For all $n = 1, 2, 3, \ldots$, each of the lattices $\mathbf{Z}^n$ and D_n contains $2(n^2 - n)$ vectors x with $\langle x, x\rangle = 2$, all equivalent under automorphisms of the lattice.
iii) The same is true of D_n^+ for each $n \equiv 0 \bmod 4$, except that $D_8^+ = E_8$ has 240 such vectors, not 112. Can you find an automorphism of E_8 mapping $e_1 + e_2$ to h? [From the existence of such an automorphism we readily deduce that $\mathrm{Aut}(E_8)$ acts transitively on all 240 vectors $x \in E_8$ such that $\langle x, x\rangle = 2$; we later give several other approaches to this result, each of which also lets us calculate the number of automorphisms.]

1.3 i) The lattice A_n in the n-dimensional Euclidean space

$$\Big\{(x_0, x_1, \ldots, x_n) \in \mathbf{R}^{n+1} : \sum_{i=0}^{n} x_i = 0\Big\}$$

is defined by

$$A_n = \Big\{(x_0, x_1, \ldots, x_n) \in \mathbf{Z}^{n+1} : \sum_{i=0}^{n} x_i = 0\Big\}. \tag{11}$$

Describe the lattice $A_2 \subset \mathbf{R}^2$ geometrically. Show that each A_n is the $\mathbf{Z}$-span of the independent vectors $e_i - e_{i-1}$ $(i = 1, \ldots, n)$. Use the Gram matrix for these generators to find $\mathrm{disc}(A_n) = n + 1$. In particular $\mathrm{disc}(A_3) = \mathrm{disc}(D_3)$; show that in fact $A_3 \cong D_3$ by finding an explicit isometry.
ii) Prove that the group A_n^*/A_n of order $n + 1$ is cyclic, generated by the coset of the vector

$$-e_0 + \frac{1}{n+1}\sum_{i=0}^{n} x_i \in A_n^*.$$

Thus for each positive factor d of $n + 1$ there exists a unique lattice $A_n^{+d} \subseteq A_n^*$ that contains A_n with index d. Thus $\mathrm{disc}(A_n^{+d}) = (n+1)/d^2$. What is the dual of

A_n^{+d}? Show that A_n^{+d} is integral if and only if $d^2|n+1$, and self-dual if and only if $d^2 = n+1$. In particular we should have $A_3^{+2} \cong \mathbf{Z}^3$; prove this by finding an explicit isomorphism. Show on the other hand that $A_8^{+3} \ncong \mathbf{Z}^8$ because there is no $x \in \mathbf{Z}_8^{+3}$ such that $\langle x, x\rangle = 1$. Thus we should have $A_8^{+3} \cong E_8$; can you find an explicit isometry?

1.4 Recall that a square matrix A is called "tridiagonal" if $A_{ij} = 0$ for all i, j such that $|i-j| > 1$. Thus vectors $v_1, \ldots, v_n$ have a tridiagonal matrix if and only if any two vectors whose indices differ by more than 1 are orthogonal. An example is our generators $v_i = e_i - e_{i-1}$ of A_n. Show more generally that a lattice L has a tridiagonal Gram matrix whose entries are all integers of absolute value at most 2 if and only if L is the direct sum of lattices each isomorphic with either $\mathbf{Z}$ or some A_n. Find a tridiagonal Gram matrix for E_8 each of whose entries is one of $0, 1, 2, 4$, with 4 occurring only once.[4]

1.5 i) The integral lattice A_7^{+2} of discriminant $(7+1)/2^2 = 2$ has the special name E_7. Show that E_7 contains no vectors x such that $\langle x, x\rangle = 1$, and 126 vectors x such that $\langle x, x\rangle = 2$.
ii) Find a self-dual lattice $L \supset E_7 \oplus E_7$, and prove that it is the unique such lattice. (Hint: start by showing that L must be contained in the dual of $E_7 \oplus E_7$.) Prove that L contains no vector x such that $\langle x, x\rangle = 1$, and thus that L is not isomorphic with any of the self-dual lattices $\mathbf{Z}^{14}$, $\mathbf{Z}^6 \oplus E_8$, and $\mathbf{Z}^2 \oplus D_{12}^+$ that we already know in 14-dimensional space.

We can now account for all the self-dual lattices tabulated in [CS2, p.49] for $n < 16$: they are $\mathbf{Z}^n$, $\mathbf{Z}^{n-8} \oplus E_8$ (for $n \geq 8$), $\mathbf{Z}^{n-12} \oplus D_{12}^+$ (for $n \geq 12$), the lattice L of exercise 1.5ii for $n = 14$, and $\mathbf{Z} \oplus L$ and A_{15}^{+4} for $n = 15$. As before, we can check that A_{15}^{+4} is distinct from the other four self-dual lattices we know in this dimension by verifying that it contains no vectors x with $\langle x, x\rangle = 1$. We later give one approach to proving the completeness of that list of self-dual lattices for $n < 16$, see Proposition 7 and Exercise 3.6 below.

[4] Iwaniec [Iw, p.176] exhibits a tridiagonal Gram matrix for E_8 with much larger entries, namely $2, 2, 4, 4, 20, 12, 4, 2$ on the main diagonal and $1, 1, 3, 5, 3, 1, 1$ next to it. Curiously this already appeared in [H2, §7], presented as a quadratic polynomial in $x_1, \ldots, x_8$, though with the coefficients $2, 2, 4, 4$ misprinted as $1, 1, 2, 2$. Perhaps Iwaniec presented this form only to silently correct Hecke's error, but that still leaves the mystery of how the tridiagonal form was obtained in the first place, since a considerably simpler one was available; even if we require each off-diagonal entry to be odd, we quickly find the example with diagonal $2, 2, 2, 2, 2, 6, 2, 4$ and off-diagonal $1, 1, 1, 1, 1, 1, 3, 1$ by computer search. The same form appears (in matrix notation, with the correct coefficients) in [Sch, p.520], where it is attributed to Minkowski. Indeed Minkowski [Min, p.77] gives this tridiagonal matrix, as well as the Gram matrix for what we now call the E_8 root system (with some sign changes), with $A_{ii} = 2$, $A_{ij} = 1$ when $|i-j| = 1$, and all other A_{ij} zero except for $A_{25} = A_{52} = -1$.
We shall meet tridiagonal Gram matrices at least once later (Exercise 5.2).

2. Theta functions of lattices

By the *norm* of a vector x in a lattice or inner-product space we mean $\langle x, x\rangle$ (and not the length $\langle x, x\rangle^{1/2}$). For any lattice $L \subset \mathbf{R}^n$ and each $k \in \mathbf{R}$, define

$$N_k(L) = \#\{v \in L \mid \langle v, v\rangle = k\}. \tag{12}$$

The set is finite; indeed for any $k_0 \in \mathbf{R}$ the sum

$$\sum_{k \le k_0} N_k(L) = \#\{v \in L \mid \langle v, v\rangle \le k_0\} \tag{13}$$

is finite because L is discrete and the subset $\{\langle x, x\rangle \le k_0\}$ in $\mathbf{R}^n$ is compact (a closed ball if $k_0 \ge 0$, empty if $k_0 < 0$). The $N_k(L)$ are important invariants of L. For example, in the context of sphere packing one is interested in the *minimal (nonzero) norm* of L, which is the smallest $k > 0$ for which $N_k(L) > 0$, and the *kissing number* of L, which is the value of $N_k(L)$ for that k.[5] The *theta function* or *theta series* Θ_L is a generating function that encodes these invariants $N_k(L)$:

$$\Theta_L(q) := \sum_{v \in L} q^{\langle v, v\rangle/2} = 1 + \sum_{k>0} N_{2k}(L) q^k. \tag{14}$$

Note that $N_0(L) = 1$; the factors of 2 are convenient as we soon see. The sum in (14) converges absolutely if $0 \le q < 1$, because the sum (13) is $O_L(k_0^{n/2})$ as $k_0 \to \infty$ (indeed it is asymptotic to $c_L k_0^{n/2}$, where c_L is $\mathrm{disc}(L)^{-1/2}$ times the volume $\pi^{n/2}/\Gamma((n/2)+1)$ of a unit sphere in $\mathbf{R}^n$. This makes it easy to justify the following derivation of the product formula for the theta function of a direct sum:

$$\begin{aligned}
\Theta_{L_1 \oplus L_2}(q) &= \sum_{(v_1, v_2) \in L_1 \oplus L_2} q^{\langle v_1+v_2, v_1+v_2\rangle/2} \\
&= \sum_{(v_1, v_2) \in L_1 \oplus L_2} q^{(\langle v_1, v_1\rangle + \langle v_2, v_2\rangle)/2} \\
&= \sum_{v_1 \in L_1} q^{\langle v_1, v_1\rangle/2} \sum_{v_2 \in L_2} q^{\langle v_2, v_2\rangle/2} \\
&= \Theta_{L_1}(q)\, \Theta_{L_2}(q).
\end{aligned} \tag{15}$$

For example,

$$\Theta_{\mathbf{Z}^n}(q) = \Theta_{\mathbf{Z}}(q)^n = \left(\sum_{m=-\infty}^{\infty} q^{m^2/2} \right)^n. \tag{16}$$

[5] The minimal norm is $(2r)^2$ where r is the radius of the largest ball in $\mathbf{R}^n$ whose translates by L have disjoint interiors. These translates constitute the sphere packing associated with L. Each of them is tangent to κ others, where κ is the kissing number.

A more remarkable identity relates the theta functions of a lattice and its dual:

Proposition 1 (functional equation for theta series)
For any lattice L in $\mathbf{R}^n$ we have

$$\Theta_{L^*}(e^{-2\pi t}) = \mathrm{disc}(L)^{1/2} t^{-n/2} \Theta_L(e^{-2\pi/t}). \tag{17}$$

for all $t > 0$.

Already the first example, with $n = 1$ and $L = L^* = \mathbf{Z}$, is surprising and important:

$$\sum_{m=-\infty}^{\infty} e^{-\pi m^2 t} = t^{-1/2} \sum_{m=-\infty}^{\infty} e^{-\pi m^2/t}. \tag{18}$$

A famous application is Riemann's proof of the analytic continuation and functional equation of the zeta function $\zeta(s) = \sum_{m=1}^{\infty} m^{-s}$: multiply $\Theta_{\mathbf{Z}}(e^{-2\pi t}) - 1$ by $t^{s/2}\, dt/t$ and integrate termwise over $0 < t < \infty$ to find

$$\begin{aligned} \int_0^\infty (\Theta_{\mathbf{Z}}(e^{-2\pi t}) - 1)\, t^{s/2} \frac{dt}{t} &= 2 \sum_{m=1}^{\infty} \int_0^\infty e^{-\pi m^2 t}\, t^{s/2} \frac{dt}{t} \\ &= 2 \sum_{m=1}^{\infty} (\pi m^2)^{-s/2} \Gamma(s/2) \\ &= 2\pi^{-s/2} \Gamma(s/2) \zeta(s) = \xi(s) \end{aligned} \tag{19}$$

for $\mathrm{Re}(s) > 0$; then split the integral as $\int_0^\infty = \int_0^1 + \int_1^\infty$ and apply (18) to $\int_0^1$ to obtain

$$\xi(s) + \frac{1}{s} + \frac{1}{1-s} = \frac{1}{2} \int_1^\infty (\Theta_{\mathbf{Z}}(e^{-2\pi t}) - 1)(t^{s/2} + t^{(1-s)/2}) \frac{dt}{t} \tag{20}$$

for $0 < \mathrm{Re}(s) < 1$. To recover the analytic continuation of $\xi(s)$ to all of $\mathbf{C}$ (with simple poles at $s = 0$ and $s = 1$), observe that the integral in (20) is an analytic function of s on all of $\mathbf{C}$, because $\Theta_{\mathbf{Z}}(e^{-2\pi t}) - 1$ decays exponentially as $t \to \infty$; the functional equation $\xi(s) = \xi(1-s)$ then follows from the symmetry of the integral under $s \leftrightarrow 1 - s$.

Applying the same definite integral to Θ_L for a general lattice L in $\mathbf{R}^n$, but using t^s instead of $t^{s/2}$, we obtain

$$\int_0^\infty (\Theta_L(e^{-2\pi t}) - 1)\, t^s \frac{dt}{t} = \pi^{-s} \Gamma(s) \zeta_L(s) = \xi_L(s), \tag{21}$$

where ζ_L is the zeta function of L, defined by

$$\zeta_L(s) := \sum_{\substack{v \in L \\ v \neq 0}} \langle v, v \rangle^{-s}, \tag{22}$$

and ξ_L is defined by the last equality in (21). Transforming the integral as before, and using the functional equation (17) relating Θ_L with Θ_{L^*}, we obtain the identity

$$\xi_{L^*}\Big(\frac{n}{2} - s\Big) = \mathrm{disc}(L)^{1/2}\xi_L(s). \tag{23}$$

See Exercises 2.3 and 2.5ii below.

For a rather more frivolous application of (18) (and one admittedly unrelated to our main topic), differentiate both sides of (18) with respect to t and set $t = 1$ to find[6]

$$\frac{d}{dt}\Theta_{\mathbf{Z}}(e^{-2\pi t})\Big|_{t=1} = -\frac{1}{4}\Theta_{\mathbf{Z}}(e^{-2\pi}),$$

whence

$$1 + 2\sum_{m=1}^{\infty} e^{-\pi m^2 t} = \Theta_{\mathbf{Z}}(e^{-2\pi}) = 8\pi \sum_{m=1}^{\infty} m^2 e^{-\pi m^2 t}.$$

Therefore $8\pi < e^{\pi} + 2$, with a rather small difference

$$e^{\pi} + 2 - 8\pi \approx (32\pi - 2)e^{-3\pi} = 0.00795+ \tag{24}$$

because $e^{-3\pi}$ is tiny. Subtracting from this the error $22 - 7\pi = 0.00885+$ in the familiar approximation $\pi \approx 22/7$ yields the striking

$$e^{\pi} - \pi = 19.999099979+,$$

which has even featured in an xkcd comic [Mu]. For a variation on this theme see Exercise 2.4 below.

Returning to our main thread, we review the proof of the functional equation for Θ_L because we shall re-use the technique several times later. The key tool is the Poisson summation formula in $\mathbf{R}^n$, which relates the sums of a Schwartz function over L and of its Fourier transform over L^*. Here a *Schwartz function* is a C^∞ function $f : \mathbf{R}^n \to \mathbf{C}$ such that f and all its partial derivatives decay as $o(\langle x, x \rangle^k)$ for all k as $\langle x, x \rangle \to \infty$. The *Fourier transform* $\hat{f} : \mathbf{R}^n \to \mathbf{C}$ is defined by

$$\hat{f}(y) = \int_{x \in \mathbf{R}^n} f(x)\, e^{2\pi i \langle x, y \rangle}\, d\mu(x), \tag{25}$$

and is a Schwartz function if f is.

[6] The next identity was certainly known to Riemann, if not to Poisson himself: an equivalent form appears in the last displayed equation before formula (2) in [Ed, p.17], where Edwards recites Riemann's derivation of a series representation for $\xi(\frac{1}{2} + it)$.

Theorem 2 (Poisson summation in $\mathbf{R}^n$)
Let L be any lattice in $\mathbf{R}^n$. Then

$$\sum_{x \in L} f(x) = (\operatorname{disc} L)^{-1/2} \sum_{y \in L^*} \hat{f}(y) \tag{26}$$

for all Schwartz functions $f : \mathbf{R}^n \to \mathbf{C}$.

Note that the $y = 0$ term in (26) is

$$(\operatorname{disc} L)^{-1/2} \hat{f}(0) = \frac{1}{\operatorname{Vol}(\mathbf{R}^n/L)} \int_{x \in \mathbf{R}^n} f(x)\, d\mu(x). \tag{27}$$

After multiplying (26) by $\operatorname{disc}(L)^{1/2}$, we can thus interpret the left-hand side as a Riemann sum approximating the integral $\hat{f}(0)$. The terms $\hat{f}(y)$ for nonzero $y \in L^*$ then measure the discrepancy between this Riemann sum and the integral.

Proof of Poisson summation: Define $F : \mathbf{R}^n \to \mathbf{C}$ by

$$F(z) = \sum_{x \in L} f(x + z). \tag{28}$$

Because f is Schwartz, the sum converges absolutely to a C^∞ function, whose value at $z = 0$ is the left-hand side of (26) Since $F(z) = F(x + z)$ for all $z \in \mathbf{R}^n$ and $x \in L$, the function descends to a C^∞ function on $\mathbf{R}^n/L$, and thus has a Fourier expansion

$$F(z) = \sum_{y \in L^*} \hat{F}(-y)\, e^{2\pi i \langle y, z \rangle}, \tag{29}$$

where

$$\hat{F}(y) = \frac{1}{\operatorname{Vol}(\mathbf{R}^n/L)} \int_{z \in \mathbf{R}^n/L} F(z)\, e^{2\pi i \langle z, y \rangle}\, d\mu(z). \tag{30}$$

Note that the vectors $y \in L^*$ are exactly those for which $e^{2\pi i \langle x, y \rangle}$ is well-defined on $\mathbf{R}^n/L$. Now choose a fundamental domain R for $\mathbf{R}^n/L$; for instance, let $v_1, \ldots, v_n$ be generators of L and set $R = \{a_1 v_1 + \cdots + a_n v_n : 0 \le a_i < 1\}$. Then we have

$$\begin{aligned}
\operatorname{Vol}(\mathbf{R}^n/L)\hat{F}(y) &= \int_{z \in R} F(z)\, e^{2\pi i \langle y, z \rangle}\, d\mu(z) \\
&= \int_{z \in R} \sum_{x \in L} f(x + z)\, e^{2\pi i \langle y, z \rangle}\, d\mu(z) \\
&= \sum_{x \in L} \int_{z \in R - x} f(z)\, e^{2\pi i \langle y, z \rangle}\, d\mu(z) \\
&= \int_{z \in \mathbf{R}^n} f(z)\, e^{2\pi i \langle y, z \rangle}\, d\mu(z) \;=\; \hat{f}(y),
\end{aligned} \tag{31}$$

where we used in the last step the fact that $\mathbf{R}^n$ is the disjoint union of the translates $R - x$ of R by lattice vectors. Thus (29) becomes

$$F(z) = \frac{1}{\mathrm{Vol}(\mathbf{R}^n/L)} \sum_{y \in L^*} \hat{f}(-y)\, e^{2\pi i \langle y,z \rangle}. \tag{32}$$

Taking $z = 0$ we obtain (26), Q.E.D.

The functional equation (17) is then the special case $f(x) = \exp(-\pi\langle x, x\rangle/t)$ of (26).

Proof of the functional equation (17) for theta series: Let $f(x) = \exp(-\pi\langle x, x\rangle/t)$ in (26). We claim that $\hat{f}(y) = t^{n/2} \exp(-\pi\langle y, y\rangle t)$. Choosing any orthogonal coordinates $(x_1, \ldots, x_n)$ for $\mathbf{R}^n$, we see that the integral (25) defining $\hat{f}(y)$ factors as

$$\prod_{j=1}^{n} \int_{-\infty}^{\infty} e^{-\pi x_j^2/t} e^{2\pi i x_j y_j}\, dx_j,$$

which reduces our claim to the case $n = 1$, which is the familiar definite integral

$$\int_{-\infty}^{\infty} e^{-\pi x^2/t}\, e^{2\pi i x y}\, dx = t^{1/2} e^{-\pi t y^2}.$$

Using these f and $\hat{f}$ in the Poisson summation formula (26) we deduce the functional equation (17), Q.E.D.

Exercises

2.1 How are the theta series and zeta function of $L\langle\alpha\rangle$ related with Θ_L and ζ_L?

2.2 Having given in (16) a formula for $\Theta_{\mathbf{Z}^n}$ in terms of $\Theta_{\mathbf{Z}}$, we find formulas for the theta functions of D_n, D_n^*, and D_n^+ in terms of $\Theta_{\mathbf{Z}}$ and two further functions[7] $\Phi_{\mathbf{Z}}, \Psi_{\mathbf{Z}}$, defined for $0 \le q < 1$ by

$$\Phi\mathbf{Z}(q) \;=\; \sum_{m=-\infty}^{\infty} (-1)^m q^{m^2/2} \;=\; 1 - 2q^{1/2} + 2q^{4/2} - 2q^{9/2} + - \cdots, \tag{33}$$

$$\Psi_{\mathbf{Z}}(q) \;=\; \sum_{m=-\infty}^{\infty} q^{(m+\frac{1}{2})^2/2} \;=\; 2\big(q^{1/8} + q^{9/8} + q^{25/8} + q^{49/8} \cdots\big). \tag{34}$$

Thus

$$\Phi\_Z(q) = \Theta_{\mathbf{Z}}(q) - 2\Theta_{\mathbf{Z}}(q^4), \qquad \Psi_{\mathbf{Z}}(q) = \Theta_{\mathbf{Z}}(q^{1/4}) - \Theta_{\mathbf{Z}}(q). \tag{35}$$

[7] The use of $\Phi_{\mathbf{Z}}, \Psi_{\mathbf{Z}}$ for these functions is an ad hoc notation. Usually these would be written as the value at $u = 1$ of Jacobi's theta functions ϑ_4, ϑ_2; likewise $\Theta_{\mathbf{Z}}$ is Jacobi's ϑ_3 evaluated at $u = 1$.

Prove that for $n = 1, 2, 3, \ldots$

$$\Theta_{D_n}(q) = \frac{1}{2}\Big(\Theta_{\mathbf{Z}}(q)^n + \Phi_{\mathbf{Z}}(q)^n\Big), \quad \Theta_{D_n^*}(q) = \Theta_{\mathbf{Z}}(q)^n + \Psi_{\mathbf{Z}}(q)^n, \tag{36}$$

and for n even

$$\Theta_{D_n^+}(q) = \frac{1}{2}\Big(\Theta_{\mathbf{Z}}(q)^n + \Psi_{\mathbf{Z}}(q)^n + \Phi_{\mathbf{Z}}(q)^n\Big), \tag{37}$$

In particular, since $D_n^+ \cong \mathbf{Z}^4$ we have Jacobi's identity

$$\Theta_{\mathbf{Z}}(q)^4 = \Psi_{\mathbf{Z}}(q)^4 + \Phi_{\mathbf{Z}}(q)^4. \tag{38}$$

2.3 Complete the derivation of the functional equation (23) relating ξ_L and ξ_{L^*}, showing along the way that ξ_L is holomorphic except for simple poles at $s = 0$ and $s = n/2$. What are the residues at these poles?

2.4 Show that $A_2^* \cong A_2\langle 1/3\rangle$. (The geometric description of A_2 from Exercise 1.3i should help.) Use this and the functional equation for Θ_{A_2} to prove

$$e^{2\pi/\sqrt{3}} > 8\sqrt{3}\,\pi - 6.$$

The two sides are not nearly as close as in (24), despite $e^{-2\pi/\sqrt{3}}$ being even smaller than $e^{-\pi}$; why?

2.5 [Sphere packing bounds from Poisson summation]

2.6 We can get further use of the formula (32) by evaluating both sides at $z \notin L$. Here are two examples for the case $n = 1$:
i) Taking $z = 1/2$, $L = L^* = \mathbf{Z}$, and $f(x) = \exp(-\pi x^2/t)$, obtain a functional equation relating $\Psi_{\mathbf{Z}}$ and $\Phi_{\mathbf{Z}}$. Check that this equation is consistent with the result of applying the $\Theta_{\mathbf{Z}}$ functional equation (18) to the formulas (35) for $\Psi_{\mathbf{Z}}$ and $\Phi_{\mathbf{Z}}$ in terms of $\Theta_{\mathbf{Z}}$. Verify also that your equation, and the formulas (36,37), are consistent with (17) for $L = D_n$ and $L = D_n^+$.
ii) Let χ_{12} be the Dirichlet character mod 12 defined by $\chi_{12}(1) = \chi_{12}(-1) = 1$, $\chi_{12}(5) = \chi_{12}(-5) = -1$ (and $\chi_{12}(m) = 0$ if $2|m$ or $3|m$). Show that if f is a Schwartz function on $\mathbf{R}$ then

$$\sum_{m=-\infty}^{\infty} \chi_{12}(m) f(m) = 12^{-1/2} \sum_{n=-\infty}^{\infty} \chi_{12}(n) \hat{f}(n/12).$$

Letting $f(x) = e^{-\pi x^2 t}$, obtain an identity analogous to (18), and deduce a functional equation for the Dirichlet L-series $L(s, \chi_{12}) = \sum_{m=1}^{\infty} \chi_{12}(m) m^{-s}$.

2.7 The Schwartz condition is much more restrictive than is needed to justify Poisson summation (though Schwartz functions suffice for all our applications to lattices and theta series).[8] Show for instance that our derivation of (26) is valid also for $n = 1$ and $f(x) = e^{-|x|}$, $\hat{f}(y) = 2/((2\pi y)^2 + 1)$, and use this to evaluate in closed form $\sum_{n=1}^{\infty} 1/(n^2 + c^2)$ for $c > 0$. Verify that your answer approaches $\zeta(2) = \pi^2/6$ as $c \to 0$.

3. Theta functions of self-dual lattices as modular forms for Γ_+

We next consider Θ_L as a function of a complex variable. For general lattices L we cannot make sense of $\Theta_L(q)$ as a function of q in a neighborhood of $q = 0$ in $\mathbf{C}$, because the exponents $\langle v, v\rangle/2$ in (14) need not be integers. However, the change of variables $q = e^{-2\pi t}$ suggested by the functional equation (17) yields a function of t that extends to a holomorphic function on the half-plane $\mathrm{Re}(t) > 0$. That functional equation then extends to this half-plane, either by analytic continuation or by using the same proof.

Suppose now that L is self-dual. Then $\Theta_L = \Theta_{L^*}$, so the functional equation (14) relates the values at t and $1/t$ of the same function. Also, each exponent $\langle v, v\rangle/2$ is in $\frac{1}{2}\mathbf{Z}$ because L is integral. Thus $\Theta_L(e^{-2\pi t})$ is also invariant under the imaginary translation $t \mapsto t + 2i$. Combining this invariance with the functional equation we then obtain further identities, one for each fractional linear transformation generated by $t \mapsto 1/t$ and $t \mapsto t + 2i$.

We make the coefficients of these transformations integral using the further change of variable $t = i\tau$. Then τ is in the Poincaré upper half-plane

$$\mathcal{H} = \{\tau \in \mathbf{C} \mid \mathrm{Im}(\tau) > 0\},$$

and

$$q = e^{2\pi i\tau}. \tag{39}$$

Our transformations $t \mapsto 1/t$ and $t \mapsto t + 2i$ then become

$$S : \tau \mapsto -1/\tau, \qquad T^2 : \tau \mapsto \tau + 2$$

acting on $\mathcal{H}$. We use the notation T^2 because we later need also the translation $T : \tau \mapsto \tau + 1$.

Recall that the group of orientation-preserving isometries of $\mathcal{H}$ with respect to the hyperbolic metric $|d\tau|/\mathrm{Im}(\tau)$ consists of the fractional linear transformations

[8] For example, Poisson summation holds if there exists $\delta > 0$ for which both $f(x)$ and $\hat{f}(x)$ are $O(\langle x, x\rangle^{-(n/2)-\delta})$ as $\langle x, x\rangle \to \infty$ [SW, Ch.VII, Cor. 2.6]; this includes the example $f(x) = e^{-|x|}$ in Exercise 2.6 with $n = 1$.

$\gamma : \tau \mapsto (a\tau + b)/(c\tau + d)$ with $a, b, c, d \in \mathbf{R}$ and $ad - bc > 0$, uniquely determined up to scaling $(a, b, c, d) \mapsto (\lambda a, \lambda b, \lambda c, \lambda d)$. We may regard γ as the projective linear transformation of the $(\tau : 1)$-line that takes $\binom{\tau}{1}$ to $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\binom{\tau}{1}$. Thus we compose γ's by multiplying the corresponding 2×2 matrices $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Requiring $ad - bc = 1$ determines λ up to sign; we thus identify our group of fractional linear transformation of $\mathcal{H}$ with $\mathrm{PSL}_2(\mathbf{R}) = \mathrm{SL}_2(\mathbf{R})/\{\pm 1\}$.

Now the maps $S : \tau \mapsto -1/\tau$ and $T : \tau \mapsto \tau + 2$ become the integer matrices $\mathsf{S} = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $\mathsf{T}^2 = \left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)$ of determinant 1. These generate a subgroup $\langle S, T^2 \rangle$ of the full modular group

$$\Gamma := \mathrm{SL}_2(\mathbf{Z})/\{\pm 1\} = \mathrm{PSL}_2(\mathbf{Z})$$

Because T^2 is congruent mod 2 to the identity matrix I, every $\gamma \in \langle S, T^2 \rangle$ is congruent to either I or $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right) \bmod 2$. (The reduction of $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{PSL}_2(\mathbf{Z})$ mod 2 is well-defined because $-\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \equiv \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \bmod 2$.) It turns out that this necessary condition is also sufficient: $\langle S, T^2 \rangle$ is the group, call it Γ_+, of all $\mathrm{PSL}_2(\mathbf{Z})$ matrices congruent mod 2 to either I or $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. See Exercise 3.1 below for one approach to the identification of $\langle S, T^2 \rangle$ with Γ_+.

We are thus led to define for any lattice L in $\mathbf{R}^n$ the function

$$\theta_L(\tau) := \Theta_L(e^{2\pi i\tau}) = 1 + \sum_{\substack{k>0 \\ N_k(L)\neq 0}} N_k(L) e^{\pi i k \tau}, \tag{40}$$

on $\mathcal{H}$. If L is self-dual then θ_L satisfies the functional equations

$$\theta_L(\tau + 2) = \theta_L(\tau), \quad \theta_L(-1/\tau) = (\tau/i)^{n/2}\theta_L(\tau), \tag{41}$$

where "$(\tau/i)^{n/2}$" is the nth power of the principal square root of τ/i (that is, the square root with positive real part). Iterating these functional equations yields for each $\gamma \in \langle S, T^2 \rangle = \Gamma_+$ an identity

$$\theta_L(\gamma(\tau)) = \epsilon_{c,d}^n (c\tau + d)^{n/2} \theta_L(\tau) \tag{42}$$

for some $\epsilon_{c,d} \in \mathbf{C}^*$ with $\epsilon_{c,d}^8 = 1$.[9] A holomorphic function $\phi : \mathcal{H} \to \mathbf{C}$ satisfying $\phi(\gamma(\tau)) = \epsilon_{c,d}^n (c\tau + d)^{n/2} \phi(\tau)$ for all $\gamma = \pm\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_+$ is a "weakly modular form of weight $n/2$ for Γ_+" (more fully, for Γ_+ and the $\epsilon_{c,d}^n$; these factors are said to form a "multiplier system of weight $n/2$" [Iw, 2.6] — a multiplier system being a

[9] We write $\epsilon_{c,d}$ rather than ϵ_γ because if some γ' has the same (c, d) as γ then $\gamma' = T^{2k}\gamma$ for some $k \in \mathbf{Z}$, so $\theta_L(\gamma(\tau)) = \theta_L(\gamma'(\tau))$ so γ and Γ' have the same ϵ; also, γ determines (c, d) only up to sign, and changing (c, d) to $(-c, -d)$ multiplies $\epsilon_{c,d}$ by $\pm i$.

system of factors in an identity such as (42) that is consistent with $\theta_L(\gamma_1(\gamma_2(\tau)) = \theta_L((\gamma_1\gamma_2)(\tau))$ for all choices of γ_1, γ_2.) We shall soon see that θ_L also satisfies the additional condition that make it modular, not just weakly modular; this will let us cite results from the theory of modular forms that for each n confine θ_L to an affine vector space of dimension $\lfloor n/8 \rfloor$. To motivate this additional condition we must first review a few further facts concerning the action of Γ and Γ_+ on $\mathcal{H}$.

The relevant results for Γ are very well known:

Theorem 3 *i)* $\Gamma = \mathrm{PSL}_2(\mathbf{Z})$ *is generated by* S *and* T.
ii) The action of Γ *on* $\mathcal{H}$ *has a fundamental domain*

$$\mathcal{F} = \{\tau \in \mathcal{H} : |\operatorname{Re}(\tau)| \le 1/2,\ |\tau| \ge 1\}. \tag{43}$$

The second assertion means that every Γ-orbit has a representative in $\mathcal{F}$, which is unique if it is in the interior of $\mathcal{F}$. Thus the images of $\mathcal{F}$ under Γ cover $\mathcal{H}$ and do not overlap except on the boundaries; see Figure 1 for a picture of part of this tiling of $\mathcal{H}$. See [Se, VII, Theorems 1 and 2] for an exposition that nicely proves both parts of Theorem 3 together. While $\mathcal{F}$ is not compact, it is closed in $\mathcal{H}$ and a sequence $\{z_j\}$ with no accumulation point in $\mathcal{F}$ must approach the "cusp" $i\infty$ in the sense that $\operatorname{Im} z_j \to \infty$.

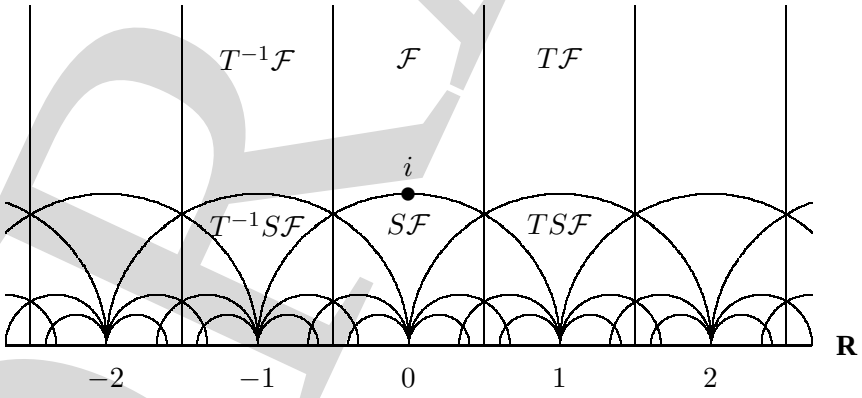

Figure 1: The fundamental domain $\mathcal{F}$ for Γ and some of its nearby images

The corresponding facts for Γ_+ are:

Theorem 4 *i)* Γ_+ *is generated by* S *and* T^2.
ii) The action of Γ_+ *on* $\mathcal{H}$ *has a fundamental domain*

$$\mathcal{F}_+ = \{\tau \in \mathcal{H} : |\operatorname{Re}(\tau)| \le 1,\ |\tau| \ge 1\}. \tag{44}$$

This can be proved by adapting the argument of [Se, VII] for Theorem 3. Alternatively it can be derived from Theorem 3 as follows. For part (i), see Exercise 3.1 below. For part (ii), we first show that $[\Gamma : \Gamma_+] = 3$. (This, as well as part (i), is noted by Serre in the concluding "Complements" section of [Se, VII].) We saw already that reduction mod 2 gives a well-defined homomorphism $\Gamma \to \mathrm{SL}_2(\mathbf{Z}/2\mathbf{Z})$. The homomorphism is readily seen to be surjective; for example, check that the images of S and T generate $\mathrm{SL}_2(\mathbf{Z}/2\mathbf{Z})$. Since $\mathrm{SL}_2(\mathbf{Z}/2\mathbf{Z})$ has order 6, the preimage Γ_+ of the 2-element subgroup $\{1, \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)\}$ has index $6/2 = 3$ as claimed. We then obtain a fundamental domain for Γ_+ by forming the union of the images of $\mathcal{F}$ under the coset representatives $1, T, TS$. The fundamental domain $\mathcal{F}_+$ is then obtained by applying $T^{-2} \in \Gamma_+$ to the right half of $T\mathcal{F} \cup TS\mathcal{F}$. See Figure 2.

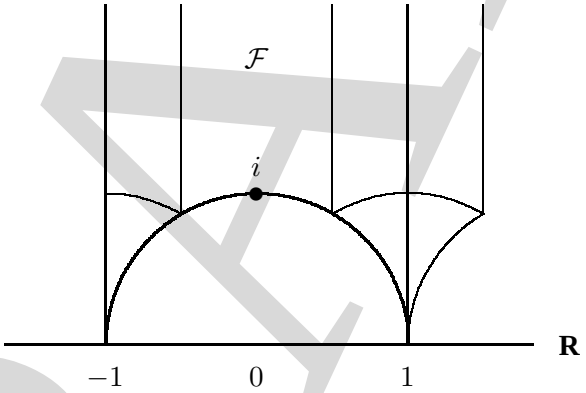

Figure 2: The fundamental domain $\mathcal{F}_+$ dissected into three images of $\mathcal{F}$

We see that $\mathcal{F}_+$ has two cusps, at ∞ and ± 1. For a general finite-index subgroup $\Gamma' \subseteq \Gamma$, we can similarly get a fundamental domain by combining $[\Gamma : \Gamma']$ images of $\mathcal{F}$, and the cusps may be identified with the orbits of the action of Γ' on $\mathbf{Q} \cup \{\infty\} = \mathbf{P}^1(\mathbf{Q})$.

Now our theta function θ_L not only satisfies the identity (42) that makes it weakly modular, but also remains bounded as $\mathrm{Im}\,\tau \to \infty$ — that is, as τ approaches the cusp $i\infty$ — because $q \to 0$ as $\tau \to i\infty$. For each $\gamma \in \Gamma_+$ it follows via (42) that $(\mathrm{Im}\,\tau)^{-n/2}\theta_L(\gamma(\tau))$ remains bounded as $\tau \to i(\infty)$, as long as τ remains in $\mathcal{F}$. Changing variables we see that $(\mathrm{Im}\,\tau)^{n/2}\theta_L(\tau)$ remains bounded as $\tau \to a/c = \gamma(i\infty)$, as long as τ remains in $\gamma\mathcal{F}$. This does not constrain the growth of $\theta_L(\tau)$ as τ approaches the cusp 1 or its images under Γ_+. But we can prove directly:

Lemma 5 *Let L be a lattice in $\mathbf{R}^n$. Then for any $t_0 > 0$ the function $\tau \mapsto (\mathrm{Im}\,\tau)^{n/2}\theta_L(\tau)$ is bounded on the strip $\{\tau \in \mathcal{H} : \mathrm{Im}\,\tau \le t_0\}$.*

Proof: If $\operatorname{Im}\tau = t$ then

$$|t^{n/2}\theta_L(\tau)| = t^{n/2}|\Theta_L(e^{2\pi i\tau})| \le t^{n/2}|\Theta_L(e^{-2\pi t})|,$$

because each of the terms $e^{\pi i\langle v,v\rangle\tau}$ in the sum defining $\Theta_L(e^{-2\pi t})$ has absolute value equal to the corresponding term $e^{-\pi\langle v,v\rangle t}$ in $\Theta_L(e^{-2\pi t})$. Hence by the functional equation (17) we have

$$|t^{n/2}\theta_L(\tau)| \le \operatorname{disc}(L)^{1/2}\Theta_L(e^{-2\pi/t}),$$

and $\Theta_L(e^{-2\pi/t}) \le \Theta_L(e^{-2\pi/t_0})$, again because the inequality holds termwise. This gives the upper bound $\operatorname{disc}(L)^{1/2}\Theta_L(e^{-2\pi/t_0})$ on $|(\operatorname{Im}\tau)^{n/2}\theta_L(\tau)|$, Q.E.D.

This suggests the following definitions: fix a finite-index subgroup Γ' of Γ and some multiplier system $\{\varepsilon_{c,d}\}$ (for all (c,d) occurring as the bottom row of some $\gamma \in \Gamma'$). A holomorphic function $\phi : \mathcal{H} \to \mathbf{C}$ satisfying

$$\phi(\gamma(\tau)) = \varepsilon_{c,d}(c\tau + d)^{n/2}\phi(\tau)$$

for all $\gamma = \pm\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma'$ is a *weakly modular form of weight* $n/2$ for Γ' and the $\varepsilon_{c,d}$; if moreover $|(\operatorname{Im}\tau)^{n/2}\phi(\tau)|$ is bounded in strips $\operatorname{Im}\tau \le t_0$ then ϕ is a *modular form of weight* $n/2$ for Γ' and the $\varepsilon_{c,d}$. The boundedness condition is equivalent to one growth condition for each cusp of $\mathcal{H}/\Gamma'$. Thus Lemma 5, combined with the preceding analysis, states that *if L is self-dual then θ_L is a modular form of weight $n/2$ for Γ_+ and the multiplier system $\{\epsilon^n_{c,d}\}$*.

Given Γ', m, and the $\varepsilon_{c,d}$, these modular forms constitute a vector space over $\mathbf{C}$, denoted by $M_{n/2}(\Gamma', \{\varepsilon_{c,d}\})$, or simply $M_{n/2}(\Gamma')$ if the $\varepsilon_{c,d}$ are known. This is a finite-dimensional vector space, and much is known about the coefficients of its elements. We shall see that for any lattice $L \subset \mathbf{R}^n$ with rational inner products the theta function θ_L is in $M_{n/2}(\Gamma', \{\varepsilon_{c,d}\})$ for some Γ' and $\varepsilon_{c,d}$. This can be used to obtain very precise statements on the $N_k(L)$, that is, on the radial distribution of lattice vectors. We shall also study the angular distribution by generalizing θ_L to "weighted theta functions" and showing and show that they, too, are modular forms. For the rest of this section and the next we illustrate this by using the structure of $M_{n/2}(\Gamma')$ for $\Gamma' = \Gamma_+$ and $\Gamma' = \Gamma$ to study self-dual lattices via their vectors' radial distribution.

If $\{\varepsilon^{(1)}_{c,d}\}$ and $\{\varepsilon^{(2)}_{c,d}\}$ are multiplier systems of weights $n_1/2$ and $n_2/2$ for Γ', then $\{\varepsilon^{(1)}_{c,d}\varepsilon^{(2)}_{c,d}\}$ is a multiplier system of weight $(n_1+n_2)/2$, and the product of modular forms in $M_{n_1/2}(\Gamma', \{\varepsilon^{(1)}_{c,d}\})$ and $M_{n_2/2}(\Gamma', \{\varepsilon^{(2)}_{c,d}\})$ is in $M_{(n_1+n_2)/2}(\Gamma', \{\varepsilon^{(1)}_{c,d}\varepsilon^{(2)}_{c,d}\})$. This happens in our setting where $\Gamma' = \Gamma_+$ and we use for each weight $n/2$ the

multipliers $\epsilon_{c,d}^n$ of (42). (Note that this is consistent with the identity (15): the change of variable $q = e^{2\pi i\tau}$ transforms that identity to $\theta_{L_1}\theta_{L_2} = \theta_{L_1\oplus L_2}$, equating a modular form of weight $(n_1+n_2)/2$ with the product of forms of weights $n_1/2$ and $n_2/2$.) This gives the direct sum

$$\mathbf{M}_+ := \bigoplus_{n=0}^{\infty} M_{n/2}(\Gamma_+) \tag{45}$$

the structure of a graded algebra. For general Γ' such algebras can be quite complicated, but our $\mathbf{M}_+$ is known to have the following simple description:

Theorem 6 *The algebra $\mathbf{M}_+$ is freely generated over $\mathbf{C}$ by the modular forms $\theta_{\mathbf{Z}}$ of weight $1/2$ and θ_{E_8} of weight 4; equivalently, by $\theta_{\mathbf{Z}}$ and*[10]

$$\Delta_+ := \frac{1}{16}(\theta_{\mathbf{Z}}^8 - \theta_{E_8}) = q^{1/2} - 8q + 28q^{3/2} - 64q^2 + - \cdots . \tag{46}$$

Thus each $M_{n/2}(\Gamma_+)$ $(n \geq 0)$ has dimension $1 + \lfloor n/8 \rfloor$ and basis

$$\{\theta_{\mathbf{Z}}^{n-8m}\Delta_+^m : m = 0, 1, 2, \ldots, \lfloor n/8 \rfloor\}. \tag{47}$$

This theorem can be proved starting from the description of $\mathcal{F}_+$ in much the same way that Serre obtains the generators of the algebra of modular forms for Γ ([Se, VII, Theorem 4]; Theorem 11 below).

Corollary. *Let L be a self-dual lattice in $\mathbf{R}^n$. Then*

$$\theta_L = \theta_{\mathbf{Z}}^n + \sum_{m=1}^{\lfloor n/8 \rfloor} c_m \theta_{\mathbf{Z}}^{n-8m}\Delta_+^m \tag{48}$$

for some constants c_m $(m = 1, 2, \ldots, \lfloor n/8 \rfloor)$.

Proof: By Theorem 6 we have $\theta_L = \sum_{m=0}^{\lfloor n/8 \rfloor} c_m \theta_{\mathbf{Z}}^{n-8m}\Delta_+^m$ for some c_m $(m = 0, 1, 2, \ldots, \lfloor n/8 \rfloor)$; so we need only show $c_0 = 1$. But c_0 is the value of the sum at $q = 0$, which is $N_0(L) = 1$ as desired.

In particular, the coefficients $N_k(L)$ of θ_L for $k = 1, 2, \ldots, \lfloor n/8 \rfloor$ determine θ_L, because we can use them to calculate the c_m iteratively. We can already use this to prove our earlier claim:

[10] The notation "Δ_+" for this form is not standard; we use it here in analogy to the weight-12 form Δ which plays a similar role in the next section.

Proposition 7 *If $n < 8$ then every self-dual lattice in $\mathbf{R}^n$ is isomorphic with $\mathbf{Z}^n$.*

Proof: Here $\lfloor n/8 \rfloor = 0$ so our Corollary determines the theta series completely: $\theta_L = \theta_{\mathbf{Z}^n}$. Comparing coefficients, we deduce $N_k(L) = N_k(\mathbf{Z}^n)$ for all k. In particular, $N_1(L) = 2n$, because $N_1(\mathbf{Z}^n) = 2n$, as may be seen either directly or by expanding (16) in powers of $q^{1/2}$. Thus $N_1(L) = 2n$, so L contains n pairs $\pm v_i$ ($1 \le i \le n$) of vectors with $\langle v_i, v_i \rangle = 1$. For $i \ne j$ we then have $|\langle v_i, v_j \rangle| < 1$ by Cauchy-Schwarz; since L is integral, it follows that $\langle v_i, v_j \rangle = 0$. That is, the v_i are orthonormal. Therefore L contains their $\mathbf{Z}$-span, call it L_0, which is isomorphic with $\mathbf{Z}^n$. But then $L = L^* \subseteq L_0^* = L_0$, so $L = L_0 \cong \mathbf{Z}^n$, Q.E.D.

See Exercises 3.5 and 3.6 for the classification of self-dual $L \subset \mathbf{R}^n$ with $n = 8$ and $9 \le n \le 15$; Exercise 3.7 for an application of the case $n = 2$ to prove Fermat's two-squares theorem, which determines $N_k(\mathbf{Z}^2)$ when k is an odd prime; and Exercise 3.8 for a different construction of $\theta_{\mathbf{Z}^2}$ that yields a formula for $N_k(\mathbf{Z}^2)$ for all k.

We conclude this section with a warning: we used the fact that $\theta_L = \theta_{\mathbf{Z}^n}$ to prove $L \cong \mathbf{Z}^2$, but in general it is possible for lattices L and L' to have the same theta series (equivalently: to satisfy $N_k(L) = N_k(L')$ for all k) without being isomorphic. We say more about this later, and give an example with self-dual lattices in $\mathbf{R}^{16}$.

Exercises

3.1 Verify that the matrices $\mathsf{S} = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $\mathsf{T} = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ satisfy $\mathsf{S}^2 = (\mathsf{ST})^3 = -I$, and thus that their images S, T in $\mathrm{PSL}_2(\mathbf{Z}) = \Gamma$ satisfy $S^2 = (ST)^3 = 1$. Use this, and the fact that $\Gamma = \langle S, T \rangle$, to prove that $\Gamma_+ = \langle S, T^2 \rangle$. [Write any $\gamma \in \Gamma_+$ as $g_1 g_2 g_3 \dots g_r$ where each $g_j \in \{S, T, T^{-1}\}$, and eliminate any occurrences of SS, $T^{-1}T$, or TT^{-1} in the product. Then, working from one side (say from the left), combine pairs with $g_j = g_{j+1} = T^{\pm 1}$ to obtain a product of factors in $\{S, T^2, T^{-2}\}$. If an impasse arises, use $(ST)^3 = I$ to replace TST or $T^{-1}ST^{-1}$ by $ST^{-1}S$ or STS respectively; note that $T^{-1}ST = T^{-2}TST$ and $TST^{-1} = T^2T^{-1}ST^{-1}$. This process either terminates or proves that $\gamma \notin \Gamma_+$ by writing γ as the product of an element of $\langle S, T^2 \rangle$ with either T or TS.]

3.2 Define a q-series

$$\begin{aligned} \eta &= \sum_{n=1}^{\infty} \chi_{12}(n) q^{n^2/24} \\ &= q^{1/24}(1 - q - q^2 + q^5 + q^7 - q^{12} - q^{15} + q^{22} + - - + \cdots), \end{aligned} \quad (49)$$

where χ_{12} is the even Dirichlet character mod 12 introduced in Exercise 2.5(ii). Use the identity proved in that Exercise to show that η satisfies, for each $\gamma =$

$\pm\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$, a functional equation

$$\eta(\gamma(\tau)) = \varepsilon_{c,d}(c\tau + d)^{1/2}\eta(\tau) \tag{50}$$

for some $\varepsilon_{c,d} \in \mathbf{C}^*$ with $\varepsilon_{c,d}^{24} = 1$. We thus say that η is a modular form of weight $1/2$ for Γ and the $\varepsilon_{c,d}$ (not just weakly modular, because η is clearly bounded at the one cusp $q \to 0$).

3.3 i) Let L be an integral lattice in $\mathbf{R}^n$ that contains a vector v of norm 1. Show that every $x \in L$ can then be written uniquely as $x' + nv$ for some $n \in \mathbf{Z}$ and $x' \in L$ with $\langle x', v\rangle = 0$, and thus that such vectors x' constitute a lattice L' in the orthogonal complement of $\mathbf{R}v$, whence $L \cong L' \oplus \mathbf{Z}$.
ii) More generally, let L be an integral lattice in $\mathbf{R}^n$, and $L_1 \subset L$ a lattice in some subspace V of dimension n_1. If L_1 is self-dual, show that $L = L_1 \oplus L'$ where $L' = L \cap V^\perp$. (Hint: given $x \in L$, use the homomorphism $L_1 \to \mathbf{R}$, $v \mapsto \langle x, v\rangle$. Part (i) is the special case $L_1 = \mathbf{Z}v$.)

3.4 Either by iterating the construction in part (i) of the previous exercise, or using part (ii) directly, prove that every integral lattice L in $\mathbf{R}^n$ can be uniquely written as $L' \oplus L_1$ where $L_1 \cong \mathbf{Z}^{n_1}$ for some integer $n_1 \in [0, n]$ and L' is a lattice in $\mathbf{R}^{n-n_1}$ that contains no vectors of norm 1. (This gives an alternative conclusion of the proof that for $n < 8$ every self-dual lattice in $\mathbf{R}^n$ is isomorphic with $\mathbf{Z}^n$; here $n_1 = n$ and L' is the zero lattice.)

3.5 Show that if L is a self-dual lattice in $\mathbf{R}^8$ then either $L \cong \mathbf{Z}^8$ or $\theta_L = \theta_{E_8}$. [In particular, in the latter case $N_2(L) = N_2(E_8) = 240$. If you know about root systems then this is all you need to deduce $L \cong E_8$, because E_8 is the only root lattice of rank at most 8 with as many as 240 vectors of norm 2. Thus L has a sublattice L_1 of finite index isomorphic with E_8, and then $[L : L_1] = 1$ because $\mathrm{disc}(E_8) = 1$, so $L = L_1$.]

3.6 More generally, show that if $0 < n < 16$ and L is a self-dual lattice in $\mathbf{R}^8$ without vectors of norm 1 then $n \geq 8$ and $\theta_L = \theta_{\mathbf{Z}}^n - 2n\theta_{\mathbf{Z}}^{n-8}\Delta_+$. In particular, $N_2(L) = 2n(23 - n)$. [Using this plus the classification of root systems we can show that E_8 is the only such lattice with $n < 12$. For $9 \leq n \leq 15$, it follows from Proposition 7 and Exercise 3.3ii that the root system cannot contain E_8; this leaves only the root systems D_{12} for $n = 12$, E_7^2 for $n = 14$, and A_{15} or $A_4 \oplus D_{11}$ for $n = 15$. In each case we then know a sublattice of finite index in L. Because this sublattice must have square discriminant, $A_4 \oplus D_{11}$ cannot occur. In each of the remaining cases the root lattice determines L uniquely as we saw in Exercises 1.2, 1.3, and 1.5 above.]

3.7 (Fermat's two-squares theorem) Let p be an odd prime. Fermat's two-squares theorem asserts that p can be represented as a sum $x_1^2 + x_2^2$ of squares of two integers

if and only if $p \equiv 1 \bmod 4$, in which case the representation is unique up to the obvious transformations $x_1 \leftrightarrow -x_1$, $x_2 \leftrightarrow -x_2$, and $x_1 \leftrightarrow x_2$. Equivalently: if $p \equiv -1 \bmod 4$ then $N_p(\mathbf{Z}^2) = 0$; and if $p \equiv +1 \bmod 4$ then $N_p(\mathbf{Z}^2) = 8$ and the dihedral group of symmetries of $\mathbf{Z}^2$ acts simply transitively on the 8 vectors with $\langle x, x\rangle = p$. We give a version of a standard proof of this theorem that concludes by invoking the case $n = 2$ of Proposition 5.

If $p = x_1^2 + x_2^2$ then $x_2 \neq 0$ so we may let $r = x_1/x_2 \in \mathbf{Z}/p\mathbf{Z}$, so that $r^2 = -1$. By Legendre's criterion, -1 is a square in $\mathbf{Z}/p\mathbf{Z}$ if and only if $(-1)^{(p-1)/2} \equiv +1 \bmod p$. This already proves the case $p \equiv -1 \bmod 4$. If $p \equiv +1 \bmod 4$ then there are two square roots of -1 in $\mathbf{Z}/p\mathbf{Z}$; let r be one of them. The vectors $(x_1, x_2) \in \mathbf{Z}^2$ such that $x_1 \equiv r x_2 \bmod p$ constitute a lattice, call it L_p. We want $x \in L_p$ such that $\langle x, x\rangle = p$. Prove that $L_p\langle 1/p\rangle$ is integral. Check that $[\mathbf{Z}^2 : L_p] = p$, and thus that $\mathrm{disc}(L_p) = p^2$. Conclude that $L_p\langle 1/p\rangle$ is self-dual. Thus $L_p\langle 1/p\rangle \cong \mathbf{Z}^2$ by Proposition 5. In particular $L_p\langle 1/p\rangle$ contains four unit vectors. Use this to complete the proof of the two-squares theorem.

3.8 (The coefficients $N_k(\mathbf{Z}^2)$ of $\theta_{\mathbf{Z}^2}$)[11] Recall that the *hyperbolic cosine* is defined by $\cosh z = \cos(iz) = (e^z + e^{-z})/2$, and the *hyperbolic secant* is $\operatorname{sech} z = 1/\cosh(z) = \sec(iz) = 2/(e^z + e^{-z})$. It is a known application of contour integration that for $t > 0$ (or even $\mathrm{Re}(t) > 0$) the Fourier transform of $\operatorname{sech}(\pi t x)$ is $t^{-1}\operatorname{sech}(\pi t^{-1} y)$. Use this to prove that

$$s_2(\tau) := \sum_{m=-\infty}^{\infty} \operatorname{sech}(m\pi i\tau) = \sum_{m=-\infty}^{\infty} \sec(m\pi\tau) \tag{52}$$

is a modular form of weight 1 for Γ_+. Since $s_2(\tau) \to 1$ as $\tau \to i\infty$ we must have

[11] The formula $s_2(\tau) = \theta_{\mathbf{Z}}(\tau)^2 = \theta_{\mathbf{Z}^2}(\tau)$ proved in this Exercise is over a century old; the proof via modular forms is also not new but its origins are harder to track down. Dickson [Di, p.235] attributes the identity

$$(1 + 2q + 2q^4 + 2q^9 + \cdots)^2 = 1 + 4\frac{q}{1-q} - 4\frac{q^3}{1-q^3} + 4\frac{q^5}{1-q^5} - \cdots \tag{51}$$

to Jacobi, in a "letter to Legendre, Sept. 9, 1828"; the right-hand side is obtained from the formula (51) for $s_2(\tau)$ by interchanging the sums over m and d. For Jacobi and others, including Dickson himself [KD], the identity arises in the theory of elliptic functions and the proof does not explicitly use the fact that s_2 and $\theta_{\mathbf{Z}}^2$ are modular. Much more recently, the development in [SS, Ch.10, §3.1, pages 297–304] closely follows the modular approach of our Exercise; this approach is also noted in [Coh, p.106, Remark (6)] though without deducing the formula for $N_k(\mathbf{Z}^2)$. But this too was probably known much earlier.

$s_2(\tau) = \theta_{\mathbf{Z}}(\tau)^2 = \theta_{\mathbf{Z}^2}(\tau)$. Then expand $s_2(\tau)$ in powers of q:

$$s_2(\tau) = 1 + 4\sum_{m=1}^{\infty} \frac{q^{m/2}}{1+q^m} = 1 + 4\sum_{m=1}^{\infty}\sum_{d=1}^{\infty} \chi_4(d) q^{dm/2}, \tag{53}$$

where χ_4 is the Dirichlet character mod 4 defined by $\chi_4(\pm 1) = \pm 1$ and $\chi_4(0) = \chi_4(2) = 0$. Deduce the formula for the number $N_k(\mathbf{Z}^2)$ of representations of an integer $k > 0$ as the sum of two squares: $N_k(\mathbf{Z}^2) = 0$ if there is a prime $p \equiv 3 \bmod 4$ such that the p-valuation $v_p(k)$ is odd; otherwise,

$$N_k(\mathbf{Z}^2) = 4 \prod_{p \equiv 1 \bmod 4} (1 + v_p(k)). \tag{54}$$

Recover Fermat's two-squares theorem as the special case that k is an odd prime.

4. The characteristic coset and shadow of a self-dual lattice

For any integral lattice L the identity

$$\langle v+w, v+w\rangle = \langle v, v\rangle + 2\langle v, w\rangle + \langle w, w\rangle \tag{55}$$

yields a homomorphism

$$L \to \mathbf{Z}/2\mathbf{Z}, \quad v \mapsto \langle v, v\rangle \bmod 2. \tag{56}$$

Suppose further that $\operatorname{disc} L$ is odd. Then $\langle \cdot, \cdot \rangle \bmod 2$ gives an isomorphism from $L/2L$ to $L\operatorname{Hom}(L, \mathbf{Z}/2\mathbf{Z})$; in particular there is a unique coset C of $2L$ in L, called the *haracteristic coset* [Se, Ch. V], that maps to the homomorphism (56). Thus $w \in C$ if and only if $\langle v, w\rangle \equiv \langle v, v\rangle \bmod 2$ for all $v \in L$. Scaling by $1/2$ yields a coset of L in $\frac{1}{2}L$ called the *shadow* of L; we denote the shadow by $s(L)$. For example, the lattice $\mathbf{Z}$ has characteristic coset $C = 2\mathbf{Z} + 1$ and shadow $s(\mathbf{Z}) = \frac{1}{2}C = \mathbf{Z} + \frac{1}{2}$, and likewise for $\mathbf{Z}\langle\alpha\rangle$ for any odd integer $\alpha > 0$. The characteristic coset and shadow are additive: if L_1 and L_2 are integral lattices of odd discriminant with characteristic cosets C_1 and C_2, then the integral lattice $L_1 \oplus L_2$ has odd discriminant and its characteristic coset and shadow are $C_1 \oplus C_2$ and $s(L_1) \oplus s(L_2)$ respectively. For example, the characteristic vectors of $\mathbf{Z}^n$ are the vectors all of whose coordinates are odd.

Now the vectors in any coset of $2L$ in L have the same norm mod 4, but for the characteristic coset we actually get congruence mod 8. Thus if $w, w' \in C$ we may write $w' = w + 2v$ for some $v \in L$, so

$$\langle w', w'\rangle = \langle w + 2v, w + 2v\rangle = \langle w, w\rangle + 4(\langle v, w\rangle + \langle v, v\rangle) \equiv \langle w, w\rangle \bmod 8.$$

Thus we have a lattice invariant $\sigma(L) \bmod 8$ which is the common residue mod 8 of the norms of all characteristic vectors. (Equivalently, all shadow vectors have quarter-integral norm congruent to $\frac{1}{4}\sigma(L) \bmod 2\mathbf{Z}$.) For example, $\sigma(\mathbf{Z}\langle\alpha\rangle) = \alpha$ because every odd square is congruent to $1 \bmod 8$. Again we have additivity: if lattices L_1, L_2 have odd discriminant then $\sigma(L_1 \oplus L_2) = \sigma(L_1) \oplus \sigma(L_2)$. For example, $\sigma(\mathbf{Z}^n) = n$. The self-dual lattices D_n^+ ($4|n$) also have characteristic vectors of norm $\equiv n \bmod 8$ (see Exercise 4.1). In fact it turns out that $\sigma(L) = n$ for any self-dual lattice $L \subset \mathbf{R}^n$. This can be proved algebraically, for example by extending σ to indefinite quadratic forms of discriminant ± 1 and studying the arithmetic of such forms as in [Se, Ch.I–V]. In our positive-definite setting the characteristic coset and shadow arise naturally when we study the transformation of θ_L under arbitrary $\gamma \in \Gamma$ (not just $\gamma \in \Gamma_+$). In the course of this study we shall give a proof of $\sigma(L) = n$ for self-dual $L \subset \mathbf{R}^n$ using theta functions.

We already know $\theta_L(\gamma(\tau))$ for all $\gamma \in \Gamma_+$. We shall determine $\theta_L(T(\tau))$ and $\theta_L(TS(\tau))$; this will suffice to obtain $\theta_L(\gamma(\tau))$ for all $\gamma \in \Gamma$ because T and TS are representatives of the two nontrivial cosets of Γ_+ in Γ.

The action of $T : \tau \mapsto \tau + 1$ is easy: this map takes $e^{\pi i \tau}$ to $-e^{\pi i \tau}$, so for any integral lattice L we have simply

$$\theta_L(T(\tau)) = \theta_L(\tau + 1) = 1 + \sum_{k=1}^{\infty} (-1)^k N_k(L) e^{\pi i k \tau}.$$

The formula for $\theta_L(TS(\tau))$ can be expressed in terms of the theta series for the shadow of L; as suggested by our notation $\Psi_{\mathbf{Z}}$ in Exercise 2.2, we define

$$\Psi_L(q) := \sum_{v \in s(L)} q^{\langle v,v\rangle/2} = 1 + \sum_{k=1}^{\infty} N_{2k}(s(L)) q^k, \tag{57}$$

$$\psi_L(\tau) := \Psi_L(e^{2\pi i \tau}) = \sum_{\substack{k>0 \\ N_k(s(L))\neq 0}} N_k(s(L)) e^{\pi i k \tau}, \tag{58}$$

where

$$N_k(s(L)) = \#\{v \in s(L) \mid \langle v, v\rangle = k\}. \tag{59}$$

Then we have:

Proposition 8 *For any self-dual lattice L in $\mathbf{R}^n$ we have*

$$\theta_L(TS(\tau)) = (\tau/i)^{n/2} \psi_L(\tau) \tag{60}$$

for all $\tau \in \mathcal{H}$.

Proof: Let w be a characteristic vector. Then

$$\theta_L(T(\tau)) = \theta_L(\tau+1) = \sum_{v\in L}(-1)^{\langle v,v\rangle}e^{\pi i\langle v,v\rangle\tau} = \sum_{v\in L} e^{\pi i(\langle v,v\rangle\tau+\langle v,w\rangle)} \tag{61}$$

because $(-1)^c = e^{\pi ic}$ for any integer c. We now apply Poisson summation to the sum. The Fourier transform of $\exp\pi i(\langle x,x\rangle\tau+\langle x,w\rangle)$ is the integral over $x\in\mathbf{R}^n$ of $\exp(\pi i(\langle x,x\rangle\tau+\langle x,w+2y\rangle))$, which is to say the value at $y+\frac{w}{2}$ of the Fourier transform of $\exp\pi i\langle x,x\rangle\tau$, which we already know is $(\tau/i)^{-n/2}\exp(-\pi i\langle x,x\rangle/\tau)$. Poisson summation then gives

$$\theta_L(T(\tau)) = (\tau/i)^{-n/2}\sum_{v\in L} e^{-\pi i\langle v+\frac{w}{2},v+\frac{w}{2}\rangle/\tau} = (\tau/i)^{-n/2}\sum_{v\in s(L)} e^{-\pi i\langle c,c\rangle/\tau} \tag{62}$$

which is $(\tau/i)^{-n/2}\psi_L(S(\tau))$; replacing τ by $S\tau$ we recover (60), Q.E.D.

Corollary. *We have $\sigma(L) = n$ for any self-dual lattice L in $\mathbf{R}^n$.*

Proof: Because $\psi_L(t+1) = e^{\pi i\sigma(L)/4}\psi_L(t)$, the claim $\sigma(L)=n$ is equivalent to

$$\psi_L(t+1) = e^{\pi in/4}\psi_L(t). \tag{63}$$

Using Proposition 8, together with $S^2 = (ST)^3 = 1$ and $\theta_L(T^2\tau) = \theta_L(\tau)$, we calculate

$$\left(\frac{t+1}{i}\right)^{n/2}\psi_L(t+1) = \theta_L(TST(t)) = \theta_L(ST^{-1}S(t))$$
$$= (T^{-1}S(t)/i)^{n/2}\theta_L(T^{-1}S(t)) = \left(\frac{i(t+1)}{t}\right)^{n/2}\theta_L(TS(t))$$

(in which we used $S^2 = (ST)^3 = \mathbf{1}$ and the invariance of θ_L under T^2, and again write $n/2$ power to mean nth power of principal square root). Comparing with (60) yields the desired identity (63), Q.E.D.

See Exercise 4.4 for an alternative approach, and Exercise 4.3 for an application to the determination of $\sigma(L)$ for lattices $L\subset\mathbf{R}^n$ of (odd) prime discriminant.

The identity (60) can be used to study the norm distribution of unimodular lattices and their shadows. For instance, we have seen that the characteristic coset of $\mathbf{Z}^n$ consists of the vectors all of whose coordinates are odd, whence it contains 2^n vectors of norm n but no vectors of norm less than n. We showed in [E1] that this latter property characterizes the $\mathbf{Z}^n$ lattice:

Theorem 9 [E1] *Let $L\subset\mathbf{R}^n$ be a self-dual lattice with no characteristic vector w such that $\langle w,w\rangle < n$. Then $L\cong\mathbf{Z}^n$.*

Proof:[12] By Theorem 6 we can write

$$\theta_L = \sum_{m=0}^{\lfloor n/8 \rfloor} c_m \theta_{\mathbf{Z}}^{n-8m} \theta_{E_8}^m \tag{64}$$

for some constants c_m. We shall see that $\psi_{E_8} = \theta_{E_8}$ (Exercise 4.4 below, and at greater length in the next section). Hence by (60) we have

$$\psi_L = \sum_{m=0}^{\lfloor n/8 \rfloor} c_m \psi_{\mathbf{Z}}^{n-8m} \theta_{E_8}^m. \tag{65}$$

We have $\psi_{\mathbf{Z}} = 2q^{1/4} + O(q^{9/4})$ as $q \to 0$, while $\theta_{E_8} = 1 + O(q^2)$. By hypothesis $\psi_L = O(q^{n/4})$. Hence $c_m = 0$ for $m > 0$, and $\theta_L = c_0 \theta_{\mathbf{Z}}^n$. Since $N_0(L) = 1$ it follows that $c_0 = 1$, so L has the same theta function as $\mathbf{Z}^n$. In particular $N_1(L) = 2n$. It follows as in Exercises 3.3 and 3.4 that $L \cong \mathbf{Z}^n$, Q.E.D.

This characterization of $\mathbf{Z}^n$ answered a question that arose in the geometry of 4-manifolds (where self-dual lattices can arise via the intersection pairing on the second homology group). See also [E2] for the determination of all self-dual $L \subset \mathbf{R}^n$ whose characteristic coset has minimal norm $n - 8$, and [CS1, RS] for the use of (60) to obtain upper bounds on the minimal nonzero norm of L.

Exercises

4.1 Prove directly that $\sigma(D_n^+) = n$ for all $n \equiv 0 \bmod 4$ by showing that $(n/2)e_i$ is a characteristic vector of D_n^+ for any $i = 1, 2, \ldots, n$. Find a characteristic vector of the self-dual lattice A_{15}^{+4} (see Exercises 1.3 and 1.5 above), and check that its norm is $\equiv 7 \bmod 8$.

4.2 i) If L is an integral lattice of odd discriminant, and $L_1 \subset L$ is a sublattice of odd index, then $\sigma(L_1) = \sigma(L)$.
ii) Let $L \subset \mathbf{R}^n$ be an integral lattice of odd discriminant, and $L_1 \subset L$ a lattice of odd discriminant in some subspace $V \subset \mathbf{R}^n$. Then $L' := L \cap V^\perp$ also has odd discriminant and $\sigma(L') = \sigma(L) - \sigma(L_1)$. [Show that $L_1 \oplus L'$ has (finite and) odd index in L.]

4.3 Let $L \subset \mathbf{R}^n$ be an integral lattice whose discriminant is an odd prime p. Suppose L^* contains a vector v^* such that $p\langle v^*, v^* \rangle \equiv -1 \bmod p$. Prove that

[12] The proof in [E1] uses an alternative, and probably preferable, approach to the key fact that such a lattice must have $\theta_L = \theta_{\mathbf{Z}^n}$, which avoids the explicit determination of $M_{n/2}(\Gamma_+)$ in Theorem 6 or indeed any explicit mention of modular forms. Instead we use that fact that $\theta_{\mathbf{Z}}$ vanishes at the cusp $\tau = \pm 1$ but nowhere in $\mathcal{H}$, deducing in effect that $\theta_L / \theta_{\mathbf{Z}^n}$ is in $M_0(\Gamma_+)$ and thus constant.

$\sigma(L) = n + 1 - p$. (Apply 4.2i to a suitable lattice containing $L \oplus \mathbf{Z}\langle p\rangle$.) What happens when L^* does not contain such v^*? [In general if L is an integral lattice with $\operatorname{disc} L = p$, and $v^* \in L^*$ is any dual vector not in L, then $p\langle v^*, v^*\rangle$ is an integer not divisible by p, and all choices of v^* yield integers with the same quadratic character mod p; thus $p\langle v^*, v^*\rangle \equiv -1 \bmod p$ is one of only two possibilities.]

4.4 Show that 0 is a characteristic vector of E_8, again consistent with $\sigma(L) = n$. Use Theorem 6 to derive (63), and thus the fact that $\sigma(L) = n$, for all self-dual lattices L from the special cases $L = \mathbf{Z}$ and $L = E_8$.

4.5 Generalize formula (60) to integral lattices of odd discriminant.

4.6 Let $L \subset \mathbf{R}^n$ be an integral lattice whose characteristic coset has minimal norm $n - 8m_0$. Show that the coefficient c_m of $\theta_{\mathbf{Z}}^{n-8m}\Delta_+^m$ in the expansion (48) of θ_L vanishes for each $m > m_0$, while c_{m_0} is $(-1)^{m_0}2^{12m_0-n}N_{n-8m_0}(C)$. In particular $(-1)^m c_{m_0} > 0$.

4.7 i) [E2] Use this to prove without using root systems that if $L \subset \mathbf{R}^n$ is a unimodular lattice for some $n < 12$ such that $N_1(L) = 0$ then $n = 8$ and $\theta_L = \theta_{E_8}$. (Use the formula for θ_L obtained in Exercise 3.6 and the fact that $N_{n-8m_0}(C)$ is an integer which is even if $n - 8m_0 \neq 0$.)

ii) Deduce further that a nonzero integral lattice with no vectors of norm 1 or 2 must have rank at least 23, and if $L \subset \mathbf{R}^{23}$ is an integral lattice with no vectors of norm 1 or 2 then

$$\theta_L = \theta_{\mathbf{Z}}^{23} - 46\theta_{\mathbf{Z}}^{15}\Delta_+ = 1 + 4600q^{3/2} + 93150q^2 + 953856q^{5/2} + \cdots .$$

[It is known that there is a unique such lattice $L \subset \mathbf{R}^{23}$, the "odd Leech lattice" or "shorter Leech lattice"; it can be obtained from the Leech lattice $\Lambda_{24} \subset \mathbf{R}^{24}$ by choosing $v_0 \in \Lambda_{24}$ of norm 4 (all such v_0 are equivalent under $\operatorname{Aut}(\Lambda_{24})$), and projecting all $v \in \Lambda_{24}$ such that $2|\langle v_0, v\rangle$ to the 23-dimensional space $(\mathbf{R}v_0)^\perp$.]

5. Even self-dual lattices and their theta functions

An important special case is an integral lattice L (with no restriction on $\operatorname{disc}(L)$) for which the homomorphism (56) is zero; that is, lattices such that $\langle v, v\rangle \in 2\mathbf{Z}$ for all $v \in L$. Such a lattice is said to be *even* an integral lattice that is not even is said to be *odd* (it contains vectors of odd norm). By (55), L is automatically integral if $\langle v, v\rangle \in 2\mathbf{Z}$ for all $v \in L$. More commonly we check that L is even by verifying that it is integral and has generators of even norm; equivalently, that it has a Gram matrix with integer entries and even diagonal entries. As with integral lattices, the direct sum of even lattices is even, as is any sublattice of an even lattice.

A nonzero vector in an even lattice must have norm at least 2. A norm-2 vector in

an integral lattice L is called a *root*. A lattice generated by its roots is known as a *root lattice*; a root lattice is even because it is integral and generated by vectors of norm 2. Such lattices are surprisingly ubiquitous, playing crucial roles in the theory of Lie groups and algebras, in the study of singularities in algebraic geometry, and elsewhere; they are also important building blocks of even lattices and other integral lattices. The direct sum of root lattices is again a root lattice, and the irreducible root lattices are precisely the lattices A_n ($n \geq 1$), D_n ($n \geq 4$), and E_n ($n = 6, 7, 8$). We have already encountered all of these except E_6, which may be defined as the sublattice of E_8 orthogonal to any isometric copy of A_2 in E_8 (all are equivalent under $\mathrm{Aut}(E_8)$). For instance, having constructed E_8 as D_8^+ we may choose the copy spanned by $e_1 - e_2$ and $e_2 - e_3$, and then E_6 is the sublattice consisting of vectors whose first three coordinates are equal. See Exercise 5.1 for the discriminants and roots of the lattices A_n, D_n, E_n. For any integral lattice L the "root (sub)lattice of L" is the sublattice generated by the roots of L.

If the integral lattice L has odd discriminant then L is even if and only if 0 is a characteristic vector, which is to say L is its own shadow. In this case $\sigma(L) = 0$, whence we have:

Theorem 10 (Schoeneberg [Sch, p.520]) *If $\mathbf{R}^n$ contains an even self-dual lattice then $n \equiv 0 \bmod 8$.*

Proof: We have seen that $\sigma(L) = n$ for every self-dual $L \subset \mathbf{R}^n$. If L is also even then $\sigma(L) = 0$ so n vanishes in $\mathbf{Z}/8\mathbf{Z}$. Q.E.D.

Theorem 10 can be used as the starting point for proving congruences mod 8 on the ranks of even lattices of other discriminants; see for example 5.4i–iii (discriminant 2) and 5.7i–ii (even lattices L such as D_4 for which $L^*/L \cong (\mathbf{Z}/2\mathbf{Z})^{n/2}$). Exercise 5.2 gives a more offbeat application of Theorem 10 to Diophantine equations.

A self-dual lattice is said to be of *Type I* or *Type II* depending on whether it is odd or even. The necessary condition $8|n$ on the existence of Type II lattice in $\mathbf{R}^n$ is also sufficient: if $8|n$ then $\mathbf{R}^n$ contains the Type II lattice $E_8^{n/8}$ (direct sum of $n/8$ copies of E_8). For $n = 8$ this is the only such lattice. For $n = 16$ there is another Type II lattice, namely D_{16}^+, which is not isomorphic with E_8^2 because E_8^2 is generated by its roots and D_{16}^+ is not. We shall soon see that these are the only Type II lattices in $\mathbf{R}^{16}$ up to isomorphism. In $\mathbf{R}^{24}$ there are 24, as first shown with great effort by Niemeier [Ni]. The classification was considerably simplified by Venkov [Ve], using the weighted theta functions we introduce in the next section. The Niemeier lattices are ubiquitous in many other classification problems for lattices of rank up to 24. See Exercise 5.4iv for an example of one

such use.

If L is a Type II lattice then its theta function θ_L is invariant under T, not just T^2, because q never appears raised to half-integral powers, only integral ones. Because θ_L still transforms under S by Proposition 1, we conclude that it transforms under the full modular group Γ generated by S and T. Moreover, since $8|n$ the S transformation simplifies to $\theta_L(S(\tau)) = \tau^{n/2}\theta_L(\tau)$, whence we have for all $\gamma = \pm\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$ the identity

$$\theta_L(\gamma(\tau)) = (c\tau + d)^{n/2}\theta_L(\tau). \tag{66}$$

In other words, *the theta function of a Type II lattice $L \in \mathbf{R}^n$ is a modular form of weight $n/2$ for Γ* with trivial multiplier system.

While the weight of θ_L is a multiple of 4, the transformation $(c\tau + d)^k\phi(\tau) = \phi(\gamma(\tau))$ is well-defined for all even k (because τ determines $\pm(c\tau + d)$); as before we say ϕ is a *modular form of weight k for Γ* if it is a holomorphic function on $\mathcal{H}$ that satisfies $\phi(\gamma(\tau)) = (c\tau + d)^k\phi(\tau)$ for all $\gamma \in \Gamma$ and ϕ remains bounded as $\tau \to i\infty$. We shall later need such forms also for $k \equiv 2 \bmod 4$, so we consider them together. Again the modular forms constitute a graded algebra, which we call

$$\mathbf{M} := \bigoplus_{\substack{k \geq 0 \\ k \text{ even}}} M_k(\Gamma); \tag{67}$$

and again the algebra turns out to be freely generated. Here these generators are the normalized Eisentein series[13]

$$\mathsf{E}_4 \;=\; 1 + 240\sum_{m=1}^{\infty} \frac{m^3 q^m}{1-q^m} = 1 + 240q + 2160q^2 + 6720q^3 + \cdots, \tag{68}$$

$$\mathsf{E}_6 \;=\; 1 - 504\sum_{m=1}^{\infty} \frac{m^5 q^m}{1-q^m} = 1 - 504q - 16632q^2 - 122976q^3 - \cdots, \tag{69}$$

of weights 4 and 6 respectively. The analogue of Theorem 6 thus reads:

Theorem 11 *The algebra $\mathbf{M}$ is freely generated over $\mathbf{C}$ by the modular forms E_4 and E_6. In other words, each $M_n(\Gamma)$ $(n = 0, 2, 4, \ldots)$ has basis*

$$\{\mathsf{E}_4^a \mathsf{E}_6^b : a, b \geq 0, 4a + 6b = n\}. \tag{70}$$

[13] Warning: there is an unfortunate but unavoidable notational collision here between the lattice E_8 and the Eisenstein series E_k. Even worse, the usual indexing of Eisenstein series makes θ_{E_8} equal E_4, not E_8 which is the theta function of each of the Type II lattices in $\mathbf{R}^{16}$. We always use a sanserif E for the Eisenstein series — which also suggests the mnemonic that this E, being thin, has half the weight.

Hence $\dim M_n(\Gamma)$ is the number of possible (a,b) in (70), which is $\lfloor n/24 \rfloor$ if $n \equiv 2 \bmod 12$ and $1 + \lfloor n/24 \rfloor$ for other even $n \geq 0$:

n	$\dim M_n$	n	$\dim M_n$	n	$\dim M_n$	$\cdots$
0	1	12	2	24	3	$\cdots$
2	0	14	1	26	2	$\cdots$
4	1	16	2	28	3	$\cdots$
6	1	18	2	30	3	$\cdots$
8	1	20	2	32	3	$\cdots$
10	1	22	2	34	3	$\cdots$

(71)

See [Se, VII, Theorem 4] for a classic exposition of the proof.

Corollary. *Let L be an even self-dual lattice in $\mathbf{R}^n$. Then*

$$\theta_L = \mathsf{E}_4^{n/8} + \sum_{m=1}^{\lfloor n/24 \rfloor} c_m \mathsf{E}_4^{n-24m} \Delta^m \tag{72}$$

for some constants c_m $(m = 1, 2, \ldots, \lfloor n/8 \rfloor)$, where

$$\Delta := \frac{\mathsf{E}_4^3 - \mathsf{E}_6^2}{12^3} = 1 - 24q + 252q^2 - 1472q^3 + 4830q^4 - 6048q^5 - 16744q^6 \cdots \tag{73}$$

Proof: Because $8|n$ any element of $M_{n/2}(\Gamma)$ is a polynomial in E_4 and E_6^2. We may thus write $\theta_L = \sum_{m=0}^{\lfloor n/24 \rfloor} c_m \theta_{\mathbf{Z}}^{n-8m} \Delta^m$ for some c_m $(m = 0, 1, 2, \ldots, \lfloor n/24 \rfloor)$, and as with (48) we use evaluate at $q = 0$ to obtain $c_0 = N_0(L) = 1$ as desired.

Again it follows that the coefficients $N_{2k}(L)$ of θ_L for $k = 1, 2, \ldots, \lfloor n/24 \rfloor$ determine θ_L, because we can use them to calculate the c_m iteratively.

In particular, when $n < 24$ there are no undetermined coefficients, so $\theta_L = \theta_{E_8}^{n/8} = 1 + 30nq + O(q^2)$, whence L has $30n$ roots. For $n = 8$ we already know that $N_2(L) = 240$, and observed that this forces $L \cong E_8$. The power series expansion (68) then determines $N_{2d}(E_8)$ for all integers $d > 0$, namely $N_{2d}(E_8) = 240 \sum_{d|n} d^3$ [Sch, p.521]. This lattice yields a sphere packing of $\mathbf{R}^8$ that is densest in its dimension among lattice packings (this was proved in 1935 by Blichfeldt [Bl], who also obtained the corresponding results for E_6 and E_7; in dimensions $n = 1, 2, 3, 4, 5$ the densest lattice packings come from the root lattices A_1, A_2, A_3, D_4, D_5). The E_8 packing is also expected to be the densest among all packings of identical spheres in $\mathbf{R}^8$, and known to be within a factor $1 + 10^{-14}$ of the maximal density ([CK], extending the computations of [CE] in the course of giving a new proof of its optimality among lattice packings).

For $n = 16$ we have $N_2(L) = 480$, and the only root systems of rank at most 16 that have 480 roots are E_8^2 and D_{16}. In the former case the root lattice already has discriminant 1, so $L \cong E_8^2$; in the latter case, $D_{16} \subset L \subset D_{16}^*$ with each inclusion having index 2; there are three such lattices, of which one is the odd lattice $\mathbf{Z}^{16}$ and the other two are isomorphic with D_{16}^+, so $L \cong D_{16}^+$. Since $\dim M_4(\Gamma) = 1$ it also follows that θ_L is the normalized Eisenstein series E_8, which yields the closed form

$$N_{2d}(E_8^2) = N_{2d}(D_{16}^{+2}) = 480 \sum_{d|n} d^7 \tag{74}$$

for the theta function coefficients.

The existence of two inequivalent lattices E_8^2 and D_{16}^+ that cannot be distinguished by their theta functions has a nice consequence in differential geometry: as Milnor [Mil] observed, the tori $\mathbf{R}^{16}/E_8^2$ and $\mathbf{R}^{16}/D_{16}^+$ are non-isometric but "isospectral", in that their Laplacians have the same eigenvalues with the same multiplicities. In the language of [Ka], we cannot hear the difference between these two 16-dimensional drums. In general, for any lattice $L \subset \mathbf{R}^n$ the Laplacian on the torus $\mathbf{R}^n/L$ has an eigenbasis consisting of functions $x \mapsto e^{2\pi i\langle v^*, x\rangle}$ for $v^* \in L^*$; the corresponding eigenvalue is $-(2\pi)^2\langle v^*, v^*\rangle$, so the tori $\mathbf{R}^n/L_1$ and $\mathbf{R}^n/L_2$ are isospectral if and only if[14] $\theta_{L_1^*} = \theta_{L_2^*}$, or equivalently (by (17)) $\theta_{L_1} = \theta_{L_2}$. Thus if $L_1 \ncong L_2$ we obtain non-isomorphic isospectral tori. Subsequently examples with $n < 16$ were discovered; it is now known that θ_L uniquely determines L for lattices $L \subset \mathbf{R}^n$ if $n \leq 3$, but already for $n = 4$ (and thus for each $n \geq 4$, using direct sums with $\mathbf{Z}\langle\alpha\rangle^{n-4}$) there is a two-dimensional family of non-isomorphic pairs $\{L_1, L_2\}$ with the same theta functions.

Once $n > 16$ we need at least one coefficient in (72). The first example is $n = 24$, when (72) gives $\theta_L = \theta_{E_8}^3 - (720 - N_2(L))\Delta$. For example, if L has no roots then

$$\theta_L = \mathsf{E}_4^3 - 720\Delta = 1 + 196560q^2 + 16773120q^3 + 398034000q^4 + \cdots . \tag{75}$$

In particular, the minimal norm of a Type II lattice in $\mathbf{R}^{24}$ is at most 4, and if it is that large then the kissing number is 196560. Leech [Le] costructed such a lattice Λ_{24}, and Conway [Con], while studying his sporadic simple group $Co_1 = \mathrm{Aut}(\Lambda_{24})/\{\pm 1\}$, showed that Λ_{24} is the unique Type II lattice of minimal norm 4. As with E_8, the Leech lattice yields a sphere packing expected to be the densest in its dimension, now known to be the densest among lattice packing (proved in

[14] Clearly if $N_k(L_1^*) = N_k(L_2^*)$ for all k then $\theta_{L_1^*} = \theta_{L_2^*}$. To show the converse, note that if $N_k(L_1^*) = N_k(L_2^*)$ does *not* hold for all k then there is a least counterexample k_0 because the set of lattice norms is discrete. But then $e^{\pi k_0 t}(\theta_{L_1^*}(it) - \theta_{L_2^*}(it)) \to (N_{k_0}(L_1^*) - N_{k_0}(L_2^*)) \neq 0$ as $t \to \infty$, so $\theta_{L_1^*} \neq \theta_{L_2^*}$.

2004 [CK], and the only example of a lattice packing of $\mathbf{R}^n$ that has been proved optimal for any $n > 8$); it is known to be within a factor $1 + \epsilon$ of the maximal density for any 24-dimensional sphere packing, here with $\epsilon = 1.65 \cdot 10^{-30}$ ([CK], again extending the computations of [CE]).

For $n = 32$ and $n = 40$ there is again one coefficient c_1 to choose. If $N_2(L) = 0$ then $c_1 = -30n$ and we find $N_4(L) = 90n(211 - 5n)$, so again the minimal norm is 4 and the kissing number is 4. Already for $n = 32$ it is known [Ki] that there are literally millions of such lattices, and thus millions of isospectral 32-dimensional tori. Each of these lattices has kissing number 146880, and we shall see that there are enumerative properties beyond the theta coefficients $N_k(L)$ shared by all these lattices' configurations of short vectors. Similarly for $n = 40$, with an even larger number of lattices, each of kissing number 39600.

For $n = 48$ there are two unknown coefficients c_1, c_2; the q and q^2 terms in θ_L vanish if and only if $(c_1, c_2) = (-1440, 125280)$, giving

$$\begin{aligned} \theta_L &= \mathsf{E}_4^6 - 1440\mathsf{E}_4^3\Delta + 125280\Delta^2 \\ &= 1 + 52416000q^3 + 39007332000q^4 + \cdots . \end{aligned} \tag{76}$$

Thus any such lattice has minimal norm 6 and kissing number 52416000. Several such lattices have been constructed (see [CS2] for more information), and the associated sphere packings are the densest known in $\mathbf{R}^{48}$. We shall see that their configurations of short vectors must satisfy many combinatorial constraints. One might hope to use these constraints to fully describe such lattices, but so far neither a full classification nor a large lower bound on the number of lattices has been achieved.

For any $n \equiv 0 \bmod 8$, we can likewise use (72) to determine the theta series of any Type II lattice in $\mathbf{R}^n$ with no nonzero vectors of norm $2m$ or less where $m = \lfloor n/24 \rfloor$. The resulting modular form is sometimes called the "extremal theta function" in $M_{n/2}(\Gamma)$. One might hope that the q^{m+1} coefficient could vanish as well, so that the minimal norm might exceed $2m + 2$; but Siegel [Si] showed that this cannot happen. See Exercise 5.8 for one approach to this result, which Siegel proved using the Lagrange-Bürmann theorem. It follows that *a Type II lattice $L \subset \mathbf{R}^n$ has minimal norm at most $2\lfloor n/24 \rfloor + 2$*, and for any such lattice θ_L is the extremal theta function in $M_{n/2}(\Gamma)$. Such L is called *extremal*. Past $n = 8$, extremal lattices yield particularly good sphere packings when $n = 24m$, when $(2\lfloor n/24 \rfloor + 2)/n$ has a local maximum. Unfortunately none are known in this case past $m = 2$; the existence of any such lattice, and in particular the $m = 3$ case (an extremal Type II lattice in $\mathbf{R}^{72}$), is a long-standing open question. They are known for all other $n \leq 88$ (subject to $8|n$ as always) and for a few other values. It is

known that for very large n (above roughly 41000) there are no extremal Type II lattices because the extremal theta series has negative the q^{m+2} coefficient [MOS].

Exercises

5.1 i) We have already seen that D_n has $2(n^2-n)$ roots (Exercise 1.2ii), and that E_7 and E_8 have respectively 126 and 240 roots. Find the 72 roots of E_6, and the n^2+n roots of A_n.
ii) We already know that $\mathrm{disc}(A_n) = n+1$ (Exercise 1.3i) and $\mathrm{disc}(D_n) = 4$ (because $[\mathbf{Z}^n : D_n] = 2$). Show that E_6 has generators $e_i - e_{i-1}$ ($i = 4, 5, 6, 7$), $e_7 + e_8$, and $h = \frac{1}{2}\sum_{i=1}^{n} e_i$, and use this to compute that $\mathrm{disc}(E_6) = 3$.
iii) More generally we can define for each $n = 2, 3, \ldots, 8$ a sublattice E_n of E_8 by requiring the first $9-n$ coordinates to be equal. This agrees with the usual E_n for $n = 6, 7, 8$. Identify the lattices E_5, E_4, E_3 with the root lattices D_5, A_4, $A_1 \oplus A_2$ respectively, and show that E_2 is a lattice with Gram matrix $\left(\begin{smallmatrix} 2 & 1 \\ 1 & 4 \end{smallmatrix}\right)$ (not a root lattice); in each case $\mathrm{disc}(E_n) = 9-n$.

5.2 [Another application of Theorem 10][15] For positive integers a, b, c, d, let M be the tridiagonal matrix

$$M = \begin{pmatrix} 2a & 1 & 0 & 0 \\ 1 & 2b & 1 & 0 \\ 0 & 1 & 2c & 1 \\ 0 & 0 & 1 & 2d \end{pmatrix}.$$

Show that M is positive definite (hint: the special case $a = b = c = d = 2$ should be familiar), and is thus the Gram matrix of an even lattice L. Note that $M \bmod p$ has rank at least 3 for any prime p, and deduce that L^*/L is cyclic. Use Theorem 10 to conclude that $\det(M)$ cannot be a square.

[This could also have been done using Proposition 7 ($\mathbf{Z}^4$ is the only self-dual lattice in $\mathbf{R}^4$), but the approach here generalizes to $n \times n$ matrices for all even $n \not\equiv 0 \bmod 8$. Only the case $n \equiv 4 \bmod 8$ is of interest: if $n \equiv 2 \bmod 4$ then $\det(M) \equiv 3 \bmod 4$ so $\det(M)$ can never be a square. Square values can be attained when $8|n$; for example, M could be a tridiagonal Gram matrix for the lattice A_8 of discriminant $9 = 3^2$.

The case $n = 4$ can also be solved more laboriously by rewriting $\det(M) = x^2$ as

[15] I do not know the original source of this problem. Henri Cohen posted it May 25, 1990 to the NMBRTHRY mailing list, writing that it is "linked to some problems of algebraic topology" and was posed to him about 15 years earlier. Cohen no longer recalled the poser's name, but was sure that he or she knew how to solve the problem. (In later communication Cohen dated the event to a graduate course taught in 1971, but still did not know who originated the problem.) I posted the solution outlined here to the mailing list, and Atkin posted a solution using quadratic reciprocity. Henri also reported that H. Lenstra solved it, but did not specify the method.

$(4ab-1)(4cd-1) = x^2 + 4ad$. We always have $\det(M) \equiv 1 \bmod 4$; numerical computation suggests that all non-square positive integers congruent to $1 \bmod 4$ occur as $\det(M)$ for some choice of a, b, c, d, with the exception of 105 and 2961. If a, b, c, d are not required to be positive then $\det(M) = 1$ can be attained by taking $a = b = c = 0$ to obtain a Gram matrix of $\mathrm{II}_{2,2}$ (the indefinite self-dual even lattice of signature $(2, 2)$).]

5.3 [Uniqueness of E_8 (again) and $\#\operatorname{Aut}(E_8)$] We use theta functions to prove that E_8 is the unique Type II lattice in $\mathbf{R}^8$ and count its automorphisms. Let L be a Type II lattice in $\mathbf{R}^8$. Thus $\theta_L = \mathsf{E}_4$. Consider the distribution of vectors of norm at most 4 among the 2^8 cosets of $2L$ in L. The zero coset contains only 0; if r, r' are roots in the same coset then $r' = \pm r$. Thus zero and the roots account for $1 + (240/2) = 121$ cosets, leaving only $2^8 - 121 = 135$. Show that each coset contains at most 16 vectors of norm 4, and if there are as many as 16 vectors then they are $\pm 2e_i$ for some orthonormal basis $\{e_i\}_{i=1}^8$ of $\mathbf{R}^8$. But the q^2 coefficient of E_4 is exactly $2160 = 16 \cdot 135$. Hence each of our 135 leftover cosets contains 16 vectors of that form. Since they are congruent mod $2L$ we deduce that L contains each of the norm 2 vectors $e_i \pm e_j$, whence $L \supset D_8$. We have seen already that this implies that L is one of the two lattices other than $\mathbf{Z}^8$ contained between D_8 and D_8^*, whence $L \cong D_8^+ = E_8$. Show that $\operatorname{Aut} D_8$ is the hyperoctahedral group of order $2^8 8!$, and deduce that E_8 has $135 \cdot 2^7 8! = 696729600$ automorphisms.

5.4 [Bimodular even lattices] Suppose $L \subset \mathbf{R}^n$ is an even lattice of discriminant 2. (Integral lattices of discriminant 2 are sometimes called "bimodular".) Let $v^* \in L^* - L$ and $\varsigma = 2\langle v^*, v^* \rangle$.
i) Show that ς is an integer whose residue mod 4 does not depend on the choice of v^*. We may thus write $\varsigma = \varsigma(L) \in \mathbf{Z}/4\mathbf{Z}$.
ii) Show that ς is odd. [Else L^* would be an integral lattice of discriminant $1/2$.] Check that $\varsigma(A_1) = 1$ and $\varsigma(E_7) = 3$.
iii) Let $L' = E_7$ or A_1 according as $\varsigma(L) = 1$ or $\varsigma(L) = 3$. Construct an even unimodular lattice $\tilde{L}$ containing $L \oplus L'$ with index 2. Conclude that $n \equiv 1 \bmod 8$ if $\varsigma(L) = 1$ while $n \equiv 7 \bmod 8$ if $\varsigma(L) = 3$.
iv) If $n = 1$ then clearly $L \cong \mathbf{Z}\langle 2 \rangle = A_1$. Show that if $n = 7$ then $L \cong E_7$; if $n = 9$ then $L \cong A_1 \oplus E_8$; while if $n = 15$ then either $L \cong E_7 \oplus E_8$ or L is the unique even lattice containing $A_1 \oplus D_{14}$ with index 2.
[You'll use several times the existence of a bijection between copies of A_1 and E_7 in E_8, namely the orthogonal slice. For $n = 17$ such lattices correspond to Niemeier lattices $\tilde{L}$ containing E_7; there are four, with root systems $D_{10} \oplus E_7^2$, $A_{17} \oplus E_7$, E_8^3, and $D_{16} \oplus E_8$. The first of these yields L with root sublattice $D_{10} \oplus E_7$, which has index 2 in L; the other three yield the lattices A_{17}^{+3}, $A_1 \oplus E_8^2$, and $A_1 \oplus D_{16}^{+2}$ respectively. For $n = 23$ we need a Niemeier lattice with a root,

and there are 31 distinct possibilities. Borcherds [Bo] enumerated 121 lattices for $n = 25$ using his analysis of the even Lorentzian lattice $\mathrm{II}_{25,1}$. For $n \geq 31$ an enumeration is likelye hopeless: by the mass formula there are millions of them for each $n = 32 \pm 1$, and many more for even larger n.]

5.5 [The theta function of a bimodular even lattice, and the Kohnen space] Let $L \subset \mathbf{R}^n$ be an even lattice of discriminant 2, and this time set $\psi_L = \theta_{L^*} - \theta_L$, the theta function of the nontrivial coset of L in L^*. Thus by the functional equation

$$\psi_L(\tau) = 2^{1/2}(\tau/i)^{n/2}\theta_L(S(\tau)) - \theta_L(\tau).$$

Also $\theta_L(T(\tau)) = \theta_L(\tau)$ as usual, and $\psi_L(T(\tau)) = i^{\varsigma(L)}\psi_L(\tau)$.
i) Use these transformation rules and the identity $(ST)^3 = 1$ to give an analytic proof that $n \equiv 1$ or $7 \bmod 8$ according as $\varsigma(L) = 1$ or $\varsigma(L) = 3$.
ii) Show that θ_L and ψ_L are modular forms of weight $n/2$ for $\Gamma(2) = \ker(\Gamma \to \mathrm{SL}_2(\mathbf{Z}/2\mathbf{Z}))$.
iii) [E_7 via "Construction A"] Define a lattice L such that $A_1^7 \subset L \subset A_1^{*7}$ as follows: identify A_1^{*7}/A_1^7 with $(\mathbf{Z}/2\mathbf{Z})^7$, and let L be the preimage of the subspace of $(\mathbf{Z}/2\mathbf{Z})^7$ consisting of zero and the cyclic permutations of $(0,0,1,0,1,1,1)$. Check that this is indeed a subspace; it is known in coding theory as the dual Hamming code. Deduce that L is a lattice, and verify that it is even and bimodular. Thus by the previous Exercise $L \cong E_7$ (can you find an explicit isomorphism?). Show that its theta function is $\theta_{\mathbf{Z}}^7 + 7\theta_{\mathbf{Z}}^3\psi_{\mathbf{Z}}^4$, and use this to compute the first few coefficients of θ_{E_7}.

5.6 [Uniqueness of E_7 (again) and $\#\operatorname{Aut}(E_7)$] Now take $n = 7$. Then $\theta_L = 1 + 126q + 756q^2 + O(q^3)$. Again we consider the distribution of vectors of norm at most 4 among the cosets of $2L$ in L. There are 2^7 cosets. The two cosets that constitute $2L^*$ contain the zero vector and no other vector of norm less than 6. Of the remaining $2^7 - 2 = 126$ cosets, $N_2(L)/2 = 63$ contain only a pair of roots, leaving 63 others. As in Exercise 5.2, each of these remaining cosets is represented by vectors $\pm 2e_i$ for some orthonormal e_i; but this time we cannot have a full orthonormal frame because then L would contain $2\mathbf{Z}^7$ with finite index, which is impossible because $2\mathbf{Z}^7$ has square discriminant and L does not. Thus there are at most 6 pairs, and since $N_4(L) = 756 = 2 \cdot 6 \cdot 63$ each of the cosets must have exactly 6. Thus $L \supset D_6$. Construct a map $L/D_6 \to D_6^*/D_6$ by sending $v \in L$ to the class of the homomorphism $D_6 \to \mathbf{Z}$, $w \mapsto \langle v, w\rangle$. Since L/D_6 is infinite cyclic and D_6^*/D_6 has exponent 2, the kernel of the map has index 1 or 2 in L/D_6. Compare discriminants to show that the index is 2 and L contains $D_6 \oplus A_1$. Finally check that there are only two even integral lattices containing $D_6 \oplus A_1$ with index 2, which are related by $\operatorname{Aut}(D_6 \oplus A_1)$. This shows that all bimodular even $L \subset \mathbf{R}^7$ are isomorphic. Since we already know such a lattice E_7 we conclude that $L \cong E_7$,

and moreover that $\#\operatorname{Aut}(E_7) = 63\#\operatorname{Aut}(D_6) = 63 \cdot 2^6 6! = 2903040$.

5.7 [2-modular lattices and their theta functions]

5.8 [Extremal theta series] Let $f(q)$ and $g(q)$ be power series of the form

$$f(q) = 1 + O(q), \quad g(q) = q + O(q^2).$$

For a nonnegative integer k and any $a_0, a_1, \ldots, a_k$ there exists a unique homogeneous polynomial $P(\cdot,\cdot)$ of degree k such that $P(f,g) = \sum_{i=0}^{k} a_i q^i + O(q^{k+1})$. [Construct this polynomial iteratively starting from the leading coefficient a_0, as we did for $(f,g) = (\theta_{\mathbf{Z}}^8, \Delta_+)$ or (E_4^3, Δ).] In particular, taking $a_0 = 1$ and $a_i = 0$ for $i = 1, 2, \ldots, k$, we find P such that $P(f,g) = 1 + Cq^{k+1} + O(q^{k+2})$. We shall show that C is $k/(k+1)$ times the $1/q$ coefficient of $f'/g^(k+1)$. In particular, $C > 0$ for all $k > 0$ if f and $1/g$ have positive coefficients. This is the case for $(f,g) = (\mathsf{E}_4^3, \Delta)$, using the expansion (68) of E_4 and the infinite product formula $\Delta = q\prod_{n=1}^{\infty}(1-q^n)^{24}$.

Divide both sides of $P(f,g) = 1 + Cq^{k+1} + O(q^{k+2})$ by g^k to obtain $p(f/g) = 1/g^k + Cq + O(q^2)$ where p is the degree-k univariate polynomial $p(X) = P(X,1)$. Since g and g/f are both of the form $q + O(q^2)$, we can find a (unique) power series $F(z) = z + b_2 z^2 + b_3 z^3 + \cdots$ such that $g = F(g/f)$. (This power series is holomorphic in a neighborhood of $z = 0$ if f and g have positive radii of convergence, but we need F only as a formal power series.) Taking $z = g/f$, we find $1/F(z)^k = p(1/z) - Cz + O(z^2)$, in which the right-hand side is the beginning of the Laurent expansion of $1/F(z)^k$ about $z = 0$. Therefore $-C$ is the z coefficient of that expansion. But then C is the residue of $-(1/F(z)^k)\,dz/z^2$ at $z = 0$.

6. Weighted theta functions

We next introduce *weighted theta functions* $\Theta_{L,P}$, $\theta_{L,P}$ of a lattice $L \subset \mathbf{R}^n$. These generalize the theta functions Θ_L, θ_L: they are generating functions that encode not just the number $N_{2k}(L)$ of lattice vectors of each norm $2k$ but also their distribution on the sphere $\langle x, x\rangle = 2k$.

Weighted theta functions are defined as follows. The *weight* P is a *harmonic polynomial* on $\mathbf{R}^n$, that is, a homogeneous polynomial whose Laplacian vanishes (we shall give a fuller description of such polynomials soon). Then

$$\Theta_{L,P}(q) \;:=\; \sum_{v\in L} P(v)\, q^{\langle v,v\rangle/2} = \sum_{k\ge 0} N_{2k}(L,P)\, q^k, \tag{77}$$

$$\theta_{L,P}(\tau) \;:=\; \sum_{v\in L} P(v)\, e^{\pi i\langle v,v\rangle\tau} = \sum_{k\ge 0} N_k(L,P)\, e^{\pi i k\tau} = \Theta_{L,P}(e^{2\pi i\tau}), \tag{78}$$

where we define

$$N_k(L,P) := \sum_{\substack{v \in L \\ \langle v,v \rangle = k}} P(v). \tag{79}$$

Let $d = \deg(P)$. If $d = 0$ then P is a constant, so $\Theta_{L,P}$ and $\theta_{L,P}$ reduce to scalar multiples of Θ_L and θ_L. If d is odd then $\Theta_{L,P} = \theta_{L,P} = 0$ because the v and $-v$ terms cancel. For even $d > 0$ we get a nontrivial generalization of Θ_L and θ_L; in this case $N_0(L,P) = 0$ so the sums over k in (77, 78) may be taken over $k > 0$.

The definitions of $\Theta_{L,P}$ and $\theta_{L,P}$ make sense for any polynomial P, harmonic or not. We require that P be harmonic so that we can generalize Proposition 1 (the functional equation (17)) to weighted theta functions. Again we shall prove the functional equation using Poisson summation; here the relevant functions on $\mathbf{R}^n$ are

$$f(x) = P(x)\, e^{-\pi \langle x,x \rangle t}. \tag{80}$$

Theorem 12 *Suppose that $t > 0$ and P is a harmonic polynomial on $\mathbf{R}^n$ of degree d, and define a function $f : \mathbf{R}^n \to \mathbf{R}$ by (80). Then the Fourier transform of f is*

$$\hat{f}(y) = i^d t^{-(\frac{n}{2}+d)} P(y)\, e^{-\pi \langle y,y \rangle / t}. \tag{81}$$

This will yield:

Proposition 13 (functional equation for weighted theta series)
For any lattice L in $\mathbf{R}^n$, and any harmonic polynomial f of degree d, we have

$$\Theta_{L^*,P}(e^{-2\pi t}) = i^d \operatorname{disc}(L)^{1/2} t^{-(n/2)-d} \Theta_L(e^{-2\pi/t}) \tag{82}$$

for all $t > 0$.

Proof (modulo Theorem 12): Apply Poisson (26) to L and the function (80), and use the formula (81) for the Fourier transform of this function. Q.E.D.

Note that since $\Theta_{L,P}$ and $\Theta_{L^*,P}$ vanish identically for odd d we can write the factor i^d as $(-1)^{d/2}$.

To prove Theorem 12, and then to use Proposition 13 to study lattices, we next review some key properties of harmonic polynomials.

Let $\mathcal{P}$ be the $\mathbf{C}$-vector space of polynomials on $\mathbf{R}^n$, and $\mathcal{P}_d$ ($d = 0, 1, 2, \ldots$) its subspace of homogeneous polynomials of degree d, so that $\mathcal{P} = \bigoplus_{d=0}^{\infty} \mathcal{P}_d$. The

Laplacian is the differential operator[16]

$$\Delta = \sum_{j=1}^{n} \frac{\partial^2}{\partial x_j^2} : \ \mathcal{C}^\infty(\mathbf{R}^n) \to \mathcal{C}^\infty(\mathbf{R}^n), \quad \mathcal{P} \to \mathcal{P}, \quad \mathcal{P}_d \to \mathcal{P}_{d-2}. \tag{83}$$

Here $x_1, \ldots, x_n$ are any orthonormal coordinates on $\mathbf{R}^n$, and $\mathcal{P}_d$ is taken to be $\{0\}$ for $d < 0$. The space of harmonic polynomials of degree d is then

$$\mathcal{P}_d^0 := \ker(\Delta : \mathcal{P}_d \to \mathcal{P}_{d-2}); \tag{84}$$

this is the degree-d homogeneous part of

$$\mathcal{P}^0 := \ker(\Delta : \mathcal{P} \to \mathcal{P}). \tag{85}$$

Examples. $\mathcal{P}_0^0$ and $\mathcal{P}_1^0$ are the spaces of constant and linear functions respectively, of dimensions 1 and n. If $n = 1$ then $\mathcal{P}_d^0 = \{0\}$ for all $d > 1$. If $n = 2$ then $\mathcal{P}_d^0$ is 2-dimensional for each $d > 0$; see Exercise 6.3 below for an explicit basis. For any n, A quadratic polynomial $P = \sum_{1 \le j \le k \le n} a_{jk} x_j x_k$ is harmonic if and only if $\sum_{j=1}^n a_{jj} = 0$, because ΔP is the constant polynomial $2 \sum_{j=1}^n a_{jj}$.

We shall see that $\Delta : \mathcal{P}_d \to \mathcal{P}_{d-2}$ is surjective, whence

$$\dim \mathcal{P}_d^0 = \dim(\mathcal{P}_d) - \dim(\mathcal{P}_{d-2}) = \binom{n+d-1}{d} - \binom{n+d-3}{d}. \tag{86}$$

Indeed we shall give a more precise result using two further operators on $\mathcal{C}^\infty(\mathbf{R}^n)$ and on its subspace $\mathcal{P}$. The first is[17]

$$\mathsf{E} := x \cdot \nabla = \sum_{j=1}^{n} x_j \frac{\partial}{\partial x_j}. \tag{87}$$

Euler proved that if $P \in \mathcal{C}^\infty(\mathbf{R}^n)$ is homogeneous of degree d then $\mathsf{E}P = d \cdot P$; in particular $\mathcal{P}_d$ is the d-eigenspace of $\mathsf{E}|_{\mathcal{P}}$. The second operator is multiplication by the norm:

$$\mathsf{F} := \langle x, x \rangle = \sum_{j=1}^{n} x_j^2, \quad P \mapsto \langle x, x \rangle P. \tag{88}$$

[16]The use of Δ for both this operator and the modular form $\eta^{24} = q \prod_{n=1}^{\infty} (1-q^n)^{24}$ may be unfortunate, but should not cause confusion because the two Δ's never appear together outside this footnote. The alternative notation L for the Laplacian would be much worse when we regularly use L for a lattice.

[17]Fortunately this operator will never appear together with an Eisenstein series E_{2k}…

Clearly F injects each $\mathcal{P}_d$ into $\mathcal{P}_{d+2}$. Thus $\mathcal{P}_d^0 = \ker(\mathsf{F}\Delta : \mathcal{P}_d \to \mathcal{P}_d)$; that is, $\mathcal{P}_d^0$ is the zero eigenspace of the operator $\mathsf{F}\Delta$ on $\mathcal{P}_d$. We next show that the other eigenspaces are $\mathsf{F}^k\mathcal{P}_{d-2k}^0$ for $k = 1, 2, \ldots, \lfloor d/2 \rfloor$, and that $\mathcal{P}_d$ is the direct sum of these eigenspaces, from which the surjectivity of $\Delta : \mathcal{P}_d \to \mathcal{P}_{d-2}$ will follow as a corollary.

We begin with by finding the commutators of $\Delta, \mathsf{E}, \mathsf{F}$. Recall that the *commutator* of any two operators A, B on some vector space is

$$[A, B] = AB - BA = -[B, A]. \tag{89}$$

For example, $[x_j, x_k] = [\partial/\partial x_j, \partial/\partial x_k] = 0$ for all j, k, while $[\partial/\partial x_j, x_k] = \delta_{jk}$ (Kronecker delta).

Lemma 14 (Commutation relations for $\Delta, \mathsf{E}, \mathsf{F}$). *We have*

$$[\Delta, \mathsf{F}] = 4\mathsf{E} + 2n, \quad [\mathsf{E}, \Delta] = -2\Delta, \quad [\mathsf{E}, \mathsf{F}] = +2\mathsf{F}. \tag{90}$$

Proof: These are direct computations using the pairwise commutators of the operators x_j and $\partial/\partial x_k$. For example, $[\partial^2/\partial x_j^2, x_k^2] = 0$ unless $j = k$, and then we calculate

$$\frac{\partial^2}{\partial x_j^2} \circ x_j^2 = \frac{\partial}{\partial x_j} \circ \Big(x_j \frac{\partial}{\partial x_j} + \big[\frac{\partial}{\partial x_j}, x_j\big]\Big) \circ x_j = \Big(\frac{\partial}{\partial x_j} \circ x_j\Big)^2 + \frac{\partial}{\partial x_j} \circ x_j$$

$$= \Big(x_j \frac{\partial}{\partial x_j} + 1\Big)^2 + x_j \frac{\partial}{\partial x_j} + 1 = \Big(x_j \frac{\partial}{\partial x_j}\Big)^2 + 3x_j \frac{\partial}{\partial x_j} + 2 = x_j^2 \frac{\partial^2}{\partial x_j^2} + 4x_j \frac{\partial}{\partial x_j} + 2,$$

whence

$$\big[\frac{\partial^2}{\partial x_j^2}, x_k^2\big] = \delta_{jk}\Big(4x_j \frac{\partial}{\partial x_j} + 2\Big).$$

Summing over $j, k = 1, \ldots, n$ yields $[\Delta, \mathsf{F}] = 4\mathsf{E} + 2n$ as claimed. We complete the proof of Lemma 14 by verifying the remaining two identities in (90) via a similar but easier calculation (see Exercise 6.1 below), Q.E.D.

For a further check on the formulas for $[\mathsf{E}, \Delta]$ and $[\mathsf{E}, \mathsf{F}]$, note that they are consistent with the action of $\Delta, \mathsf{E}, \mathsf{F}$ on the $\mathcal{P}_d$: if $P \in \mathcal{P}_d$ then $\Delta P \in \mathcal{P}_{d-2}$ yields

$$[\mathsf{E}, \Delta]P = \mathsf{E}\Delta P - \Delta \mathsf{E} P = (d-2)\Delta P - \Delta(d \cdot P) = -2\Delta P,$$

consistent with $[\mathsf{E}, \Delta] = -2\Delta$, and likewise for the third part $[\mathsf{E}, \mathsf{F}] = +2\mathsf{F}$ of (90).

See the further Remarks at the end of this section for the interpretation of the commutation relations (90) (and Lemmas 17,18(ii) below) in terms of $\mathfrak{sl}_2$ and other Lie algebras and groups.

Now suppose $P \in \mathcal{P}_d$ is in the λ-eigenspace of $\mathsf{F}\Delta$ for some λ. Then $\langle x, x\rangle P = \mathsf{F}P$ is in the $(\lambda + 4d + 2n)$-eigenspace of $\mathsf{F}\Delta$ acting on $\mathcal{P}_{d+2}$, because

$$\mathsf{F}\Delta\mathsf{F}P = \mathsf{F}(\mathsf{F}\Delta + [\Delta, \mathsf{F}])P = \mathsf{F}(\mathsf{F}\Delta + 4\mathsf{E} + 2n)P = \mathsf{F}(\lambda + 4d + 2n)P.$$

By induction on $k = 0, 1, 2, \ldots$ it follows that $\mathsf{F}^k P$ is an eigenvector of $\mathsf{F}\Delta|_{\mathcal{P}_{d+2k}}$ with eigenvalue

$$\lambda + \sum_{j=0}^{k-1} 4(d + 2j) + 2n = \lambda + k\big(4(d + k - 1) + 2n\big).$$

Replacing d by $d - 2k$ and taking $\lambda = 0$, we see that if $P \in \mathcal{P}^0_{d-2k}$ then $\mathsf{F}^k P$ is an eigenvector of $\mathsf{F}\Delta|_{\mathcal{P}_d}$ with eigenvalue

$$\lambda_d(k) := k\,(4(d - k - 1) + 2n)\,. \tag{91}$$

We next prove that this accounts for all the eigenspaces of $\mathsf{F}\Delta|_{\mathcal{P}_d}$.

Lemma 15 *Fix $d \geq 0$. For integers k, k' such that $0 \leq k < k' \leq d/2$ we have $\lambda_d(k) < \lambda_d(k')$.*

Proof: By induction it is enough to check this for $k' = k + 1$. We compute

$$\lambda_d(k + 1) - \lambda_d(k) = 2n + 4(d - 2k') \geq 2n > 0,$$

Q.E.D.

Corollary. *The sum of the subspaces $\mathsf{F}^k\mathcal{P}^0_{d-2k}$ of $\mathcal{P}_d$ over $k = 0, 1, \ldots, \lfloor d/2\rfloor$ is direct.*

Proof: Since $\mathsf{F}^k\mathcal{P}^0_{d-2k}$ is a subspace of the $\lambda_d(k)$ eigenspace of $\mathsf{F}\Delta$, it is enough to prove that the $\lambda_d(k)$ are pairwise distinct. By Lemma 15, they are strictly increasing, Q.E.D.

Proposition 16 *For $k = 0, 1, \ldots, \lfloor d/2\rfloor$, let $\mathcal{P}^k_d = \mathsf{F}^k\mathcal{P}^0_{d-2k}$. Then:*
i) The map $\Delta : \mathcal{P}_d \to \mathcal{P}_{d-2}$ is surjective.
ii) $\mathcal{P}_d = \bigoplus_{k=0}^{\lfloor d/2\rfloor} \mathcal{P}^k_d = \mathcal{P}^0_d \oplus \mathsf{F}\mathcal{P}_{d-2}$, and $\mathcal{P} = \bigoplus_{k=0}^{\infty} \mathsf{F}^k\mathcal{P}^0$.
iii) $\mathcal{P}^k_d$ is the entire $\lambda_d(k)$ eigenspace of $\mathsf{F}\Delta|_{\mathcal{P}_d}$, and $\mathsf{F}\Delta|_{\mathcal{P}_d}$ has no eigenvalues other than the $\lambda_d(k)$ for $k = 0, 1, \ldots, \lfloor d/2\rfloor$.
iv) $\dim \mathcal{P}^0_d = \dim(\mathcal{P}_d) - \dim(\mathcal{P}_{d-2})$ *as claimed in (86).*

Proof: The sum in (ii) is direct by the previous Corollary. We prove that it equals $\mathcal{P}_d$ by comparing dimensions. Since F is injective we have $\dim \mathcal{P}_d^k = \dim \mathcal{P}_{d-2k}^0$; moreover

$$\dim \mathcal{P}_{d-2k}^0 \geq \dim \mathcal{P}_{d-2k} - \dim \mathcal{P}_{d-2k-2},$$

with equality if and only if $\Delta : \mathcal{P}_{d-2k} \to \mathcal{P}_{d-2k-2}$ is surjectve, because $\mathcal{P}_{d-2k}^0 = \ker(\Delta : \mathcal{P}_{d-2k} \to \mathcal{P}_{d-2k-2})$. Hence $\dim \bigoplus_{k=0}^{\lfloor d/2 \rfloor} \mathcal{P}_d^k$ is

$$\sum_{k=0}^{\lfloor d/2 \rfloor} \dim \mathcal{P}_d^k = \sum_{k=0}^{\lfloor d/2 \rfloor} \dim \mathcal{P}_{d-2k}^0 \geq \sum_{k=0}^{\lfloor d/2 \rfloor} \dim \mathcal{P}_{d-2k} - \dim \mathcal{P}_{d-2k-2}, \qquad (92)$$

and the last sum telescopes to $\dim \mathcal{P}_d$. Thus equality holds in the last step of (92) and $\dim \bigoplus_{k=0}^{\lfloor d/2 \rfloor} \mathcal{P}_d^k = \dim \mathcal{P}_d$. The first of these proves (i) (by taking $k = 0$). The second yields

$$\mathcal{P}_d = \bigoplus_{k=0}^{\lfloor d/2 \rfloor} \mathcal{P}_d^k, \qquad (93)$$

as claimed in (ii); taking the direct sum over d yields $\mathcal{P} = \bigoplus_{k=0}^{\infty} \mathsf{F}^k \mathcal{P}^0$, also claimed in (ii). to complete the proof of (ii) we compare the decompositions (93) of $\mathcal{P}_d$ and $\mathcal{P}_{d-2}$ and note that $\mathcal{P}_d^k = \mathsf{F}\mathcal{P}_{d-2}^{k-1}$ for each $k > 0$. Claim (iii) follows because the decomposition (93) diagonalizes $\mathsf{F}\Delta|_{\mathcal{P}_d}$. Finally (iv) is the case $k = 0$ of equality in (92), Q.E.D.

Remark: Part (ii) implies $\mathcal{P}_d^0 \cap \mathsf{F}\mathcal{P}_{d-2} = \{0\}$, and thus that $\mathcal{P}^0$ contains no nonzero multiple of $\langle x, x \rangle$. Proving this was set as problem B-5 on the 2005 Putnam exam [KAL, p.736], which was the hardest of the 12 problems that year, solved by only five of the top 200 scorers [KAL, p.741]. The solution printed in [KAL, p.742] uses some of the ingredients used here to prove Proposition 16.[18]

We next characterize the functions $f(x) = P(x)\, e^{-\pi \langle x,x \rangle t}$ of (80) using the operators $\Delta, \mathsf{E}, \mathsf{F}$. For $t \in \mathbf{C}$ define an operator

$$G_t : \mathcal{C}^\infty(\mathbf{R}^n) \to \mathcal{C}^\infty(\mathbf{R}^n), \quad g \mapsto e^{-\pi t \langle x,x \rangle} g \qquad (94)$$

[18] Suppose $\langle x, x \rangle | P$ and $\Delta P = 0$. Write $P = \sum_{d \geq 0} P_d$ with each $P_d \in \mathcal{P}_d$. Then $\langle x, x \rangle | P_d$ and $\Delta P_d = 0$ for each d, and we may choose d so that $P_d \neq 0$. Let m be the largest integer such that $P_d = \langle x, x \rangle^m Q$ for some polynomial Q; by assumption $m > 0$. In our notations, then, $Q \in \mathcal{P}_{d-2m}$ with $\Delta \mathsf{F}^m Q = 0$ and $Q \notin \mathsf{F}\mathcal{P}$. Using in effect the formula for $[\Delta, \mathsf{F}]$ and Euler's description of E, compute that

$$\Delta \mathsf{F}^m Q = \mathsf{F}^{m-1}[\mathsf{F}\Delta + 2m(n + 2(d - m - 1))]Q$$

for all $Q \in \mathcal{P}_{d-2m}$. Thus $\Delta \mathsf{F}^m Q = 0$ implies $\mathsf{F}\Delta Q = -2m(n + 2(d - m - 1))Q$, and the factor $n + 2(d - m - 1)$ is positive (because $d \geq 2m \geq m + 1$), so $Q \in \mathsf{F}\mathcal{P}$, contradiction.

that multiplies every $\mathcal{C}^\infty$ function by the Gaussian $e^{-\pi t\langle x,x\rangle}$; these operators constitute a one-parameter group: $G_t G_{t'} = G_{t+t'}$ for all t, t'. We are then interested in $f = G_t P$ for $P \in \mathcal{P}$ in the intersection of the kernel of Δ with an eigenspace of E. If $P \in \mathcal{P}_d$ then

$$d \cdot f = G_t(d \cdot P) = G_t \mathsf{E} P = (G_t \mathsf{E} G_{-t}) G_t \mathsf{E} P = (G_t \mathsf{E} G_{-t}) f, \qquad (95)$$

so f is in the d-eigenspace of $G_t \mathsf{E} G_{-t}$; likewise $f \in \ker G_t \Delta G_{-t}$. Since our one-parameter group $\{G_t\}$ has infinitesimal generator $-\pi\mathsf{F}$, we expect that conjugation by G_t will take Δ, E to some linear combination of $\Delta, \mathsf{E}, \mathsf{F}$. Indeed we find:

Lemma 17 (Conjugation of $\Delta, \mathsf{E}, \mathsf{F}$ by G_t) *The operators G_t commute with F, and we have*

$$G_t \mathsf{E} G_{-t} = \mathsf{E} + 2\pi t \mathsf{F}, \quad G_t \Delta G_{-t} = \Delta + \pi t(4\mathsf{E} + 2n) + (2\pi t)^2 \mathsf{F}. \qquad (96)$$

Proof: As with Lemma 14, this comes down to an exercise in differential calculus. Here we start from the fact that G_t commutes with each x_j while $G_t(\partial/\partial x_j)G_{-t} = 2\pi t x_j$, whence the first formula in (96) quickly follows, while $G_t \mathsf{F} = \mathsf{F} G_t$ is immediate. A somewhat longer computation (see Exercise 6.2 below) establishes the second formula, Q.E.D.

Corollary. *The operators $\Delta, \mathsf{E}, \mathsf{F}$ act on $G_t\mathcal{P}$, and the subspace $G_t\mathcal{P}_d^0$ is the intersection of $\ker(\Delta + \pi t(4\mathsf{E} + 2n) + (2\pi t)^2\mathsf{F})$ with the d-eigenspace of $\mathsf{E} + 2\pi t\mathsf{F}$ in $G_t\mathcal{P}$.*

We next relate the Fourier transform of a Schwartz function f with the Fourier transforms of its images under $\Delta, \mathsf{E}, \mathsf{F}$, and use this to prove Theorem 12.

Lemma 18 (Conjugation of $\Delta, \mathsf{E}, \mathsf{F}$ by the Fourier transform). *Let $f : \mathbf{R}^n \to \mathbf{C}$ be any Schwartz function. Then:*
i) For each $j = 1, \ldots, n$, the Fourier transform of $x_j f$ is $(2\pi i)^{-1}\partial\hat{f}/\partial y_j$, and the Fourier transform of $\partial f/\partial x_j$ is $-2\pi i y_j \hat{f}$.
ii) The Fourier transforms of Δf, $(2\mathsf{E} + n)f$, and $\mathsf{F}f$ are respectively $-(2\pi)^2\mathsf{F}\hat{f}$, $-(2\mathsf{E} + n)\hat{f}$, and $-(2\pi)^{-2}\Delta\hat{f}$.

Proof: Again a calculus exercise, this time with definite integrals. The formula for the Fourier transform of $\partial f/\partial x_j$ is obtained by integrating by parts with respect to x_j. The Fourier transform of $x_j f$ can be obtained from this using Fourier inversion, or directly by differentiation with respect to y_j of the integral (25) that defines $\hat{f}(y)$. We then obtain (ii) by iterating the formulas in (i) to find the Fourier

transform of $\partial^2 f/\partial x_j^2$, $x_j \partial f/\partial x_j$, or $x_j^2 f$, and summing over j. The case of $\mathsf{E}f$ can be explained by writing the operator $2\mathsf{E}+n$ as $\sum_{j=1}^n x_j(\partial/\partial x_j)+(\partial/\partial x_j)\circ x_j$, Q.E.D.

We first use this to show that if $f \in G_t\mathcal{P}$ then $\hat{f} \in G_{1/t}\mathcal{P}$, that is, that $\hat{f}$ is *some* polynomial multiplied by $e^{-\pi\langle y,y\rangle/t}$; more precisely:

Proposition 19 *Let $t \in \mathbf{C}$ with $\mathrm{Re}(t) > 0$. If $f = G_t P$ for some $P \in \mathcal{P}_d$ then $\hat{f} = G_{1/t}\hat{P}$ for some $\hat{P} = \sum_{d'=0}^d \hat{P}_{d'}$ with each $\hat{P}_{d'} \in \mathcal{P}_{d'}$ and $\hat{P}_d = i^d t^{-(\frac{n}{2}+d)} P$. As before $t^{-(\frac{n}{2}+d)}$ denotes the $-(n+2d)$ power of the principal square root of t.*

Proof: We use induction on d. The base case $d = 0$ is the fact that the Fourier transform of $e^{-\pi t\langle x,x\rangle}$ is $t^{-n/2}e^{-\pi\langle y,y\rangle/t}$, which we showed already. Suppose we have established the claim for $P \in \mathcal{P}_d$. By linearity and the fact that $\mathcal{P}_{d+1}$ is spanned by its subspaces $x_j\mathcal{P}_d$, it is enough to prove the Proposition with P replaced by $x_j P$. By part (i) of Lemma 18, the Fourier transform of $G_t x_j P = x_j G_t P$ is

$$\frac{1}{2\pi i}\frac{\partial}{\partial y_j}(G_{1/t}\hat{P}) = \frac{1}{2\pi i} G_{1/t}\Big(\frac{\partial \hat{P}}{\partial y_j} - \frac{2\pi}{t} y_j \hat{P}\Big). \tag{97}$$

By the inductive hypothesis $\hat{P}$ has degree d and leading part $\hat{P}_d = i^d t^{-(\frac{n}{2}+d)} P$. Therefore the right-hand side of (97) has degree $d+1$ and leading part

$$\frac{-2\pi t^{-1}}{2\pi i}\hat{P}_d = \frac{i}{t}\hat{P}_d = i^{d+1} t^{-(\frac{n}{2}+d+1)} y_j P.$$

This completes the induction step and the proof, Q.E.D.

Proof of Theorem 12: Suppose $P \in \mathcal{P}_d^0$ and $f(x) = P(x)\, e^{-\pi\langle x,x\rangle t} = G_t P$. By the Corollary to Lemma 17,

$$(\Delta + \pi t(4\mathsf{E}+2n) + (2\pi t)^2\mathsf{F})f = 0, \qquad (\mathsf{E} + 2\pi t\mathsf{F})f = d\cdot f. \tag{98}$$

Taking the Fourier transform and applying Lemma 18(ii), we deduce

$$(-(2\pi)^2\mathsf{F} - \pi t(4\mathsf{E}+2n) - t^2\Delta)\hat{f} = 0, \quad -(\mathsf{E} + n + \frac{t}{2\pi}\Delta)\hat{f} = d\cdot\hat{f}. \tag{99}$$

Eliminating $\Delta\hat{f}$, we find $d\cdot\hat{f} = (\mathsf{E} + \frac{2\pi}{t}\mathsf{F})\hat{f}$; that is, $\hat{f}$ is in the d-eigenspace of $\mathsf{E} + 2\pi t^{-1}\mathsf{F}$. By Proposition 19 we know $\hat{f} = G_{1/t}\hat{P}$ for some $\hat{P} \in \mathcal{P}$. By Lemma 17, then, $\hat{P}$ is in the d-eigenspace of E; that is, $\hat{P} \in \mathcal{P}_d$. By Proposition 19 we conclude that $\hat{P} = i^d t^{-(\frac{n}{2}+d)} P$, Q.E.D.

Remark: Multiplying the first equation in (99) by $-t^{-2}$, we recover the first equation of (98) with t replaced by t^{-1}; this lets us show without Euler's theorem that the Fourier transform takes $G_t\mathcal{P}^0$ to $G_{1/t}\mathcal{P}^0$. See the further Remarks at the end of this section for a sketch of an alternative approach to Theorem 12, using the connection with $\mathfrak{sl}_2$ and SL_2 and avoiding Proposition 19.

A natural application of weighted theta functions is to the question of equidistribution of lattice points in spherical shells. The $N_k(L)$ lattice vectors $v \in L$ on the sphere $\{x \in \mathbf{R}^n \mid \langle x, x\rangle = k\}$ yield a configuration, call it

$$S_k(L) := k^{-1/2}\{v \in L \mid \langle v, v\rangle = k\}, \tag{100}$$

of $N_k(L)$ vectors on the unit sphere $\Sigma \subset \mathbf{R}^n$. As $k \to \infty$ through the lattice norms, are the S_k asymptotically equidistributed on Σ? Recall that a sequence $\{C_m\}_{m=1}^{\infty}$ of nonempty finite subsets of Σ are *asymptotically equidistributed* if, for every continuous function $\varphi : \Sigma \to \mathbf{C}$, the average of φ over C_m approaches the average of φ over Σ as $m \to \infty$:

$$\int_{x\in\Sigma} \varphi(x)\, d\nu_x = \lim_{m\to\infty} \frac{1}{\#C_m} \sum_{x\in C_m} \varphi(x). \tag{101}$$

Here $d\nu$ is the invariant measure on Σ such that $\int_{x\in\Sigma} d\mu_x = 1$; for instance we may define

$$\int_{x\in\Sigma} \varphi(x)\, d\nu_x = t^{n/2} \int_{x\in\mathbf{R}^n} e^{-\pi t\langle x,x\rangle}\, \varphi\left(\frac{x}{\langle x, x\rangle^{1/2}}\right) d\mu_x \tag{102}$$

for any $t > 0$. The coefficients of weighted theta functions $\Theta_{L,P}$ give us the sum in (101) when P is a harmonic polynomial. Using the decomposition (93) from Proposition 16(ii), we show that these are enough to test equidistribution:

Proposition 20 *$\mathcal{P}^0|_\Sigma$ is dense in $\mathcal{C}(\Sigma)$; that is, for every continuous $\varphi : \Sigma \to \mathbf{C}$ and any $\epsilon > 0$ there exists $P \in \mathcal{P}^0$ such that*

$$\forall x \in \Sigma : |P(x) - \varphi(x)| < \epsilon. \tag{103}$$

Proof: By the Stone–Weierstrass theorem there exists $P \in \mathcal{P}$ satisfying (103). It is thus enough to prove that for every $P \in \mathcal{P}$ there exists $Q \in \mathcal{P}^0$ such that $P = Q$ on Σ. Applying Proposition 16(ii) to each homogeneous part of P we write

$$P = \sum_{k=0}^{\lfloor \deg(P)/2\rfloor} \langle x, x\rangle^k Q_k \tag{104}$$

for some $Q_k \in \mathcal{P}^0$. Since $\langle x, x\rangle = 1$ on Σ, the polynomial $Q = \sum_k Q_k \in \mathcal{P}^0$ agrees with P on Σ, Q.E.D.

Theorem 21 *A sequence $\{C_m\}_{m=1}^{\infty}$ of nonempty finite subsets of Σ is asymptotically equidistributed if and only if*

$$\lim_{m\to\infty} \frac{1}{\#C_m} \sum_{x\in C_m} P(x) = 0 \tag{105}$$

for all harmonic polynomials P of positive degree.

Proof: For the "only if" direction, assume (101) holds for all $\varphi \in \mathcal{C}(\Sigma)$, and take $\varphi = P|_\Sigma$. We claim (101) is then equivalent to (105), i.e., that $\int_{x\in\Sigma} P(x) d\nu_x = 0$. Equivalently (see Exercise 6.5), we claim $\int_{x\in\mathbf{R}^n} P(x)\, e^{-\pi t\langle x,x\rangle}\, d\mu_x = 0$. But the integral is the value at $y = 0$ of the Fourier transform of $P(x)\, e^{-\pi t\langle x,x\rangle}$. We obtained this Fourier transform in Theorem 12; it vanishes at $y = 0$ as claimed, because $P(0) = 0$ for P of positive degree.

Since both sides of (101) are linear, we have thus proved the converse implication for functions φ that are the restriction to Σ of any finite linear combination of harmonic polynomials of positive degree. We can drop the condition of positive degree, because (101) holds automatically for $\varphi = 1$: its left-hand side equals 1, and the right-hand side reduces to $\lim_{m\to\infty} 1$. Thus (105) implies (101) for all φ of the form $P|_\Sigma$ with $P \in \mathcal{P}^0$. By Proposition 20, every continuous $\varphi : \Sigma \to \mathbf{C}$ can be uniformly approximated by such $P|_\Sigma$. Hence (101) holds for all $\varphi \in \mathcal{C}(\Sigma)$ by the following standard argument. Changing φ to P moves both $\int_{x\in\Sigma} \varphi(x)\, d\nu_x$ and each average $(\#C_m)^{-1} \sum_{x\in C_m} \varphi(x)$ by at most ϵ. For large enough m, the average of P over C_m is within ϵ of $\int_{x\in\Sigma} P(x)\, d\nu_x$. Therefore the average and integral of φ are within 3ϵ of each other. Since ϵ is arbitrary, we are done. Q.E.D.

[...]

Further *Remarks*:[19] The commutation relations in Lemma 14 are tantamount to an isomorphism of Lie algebras from $\mathfrak{sl}_2$ to the span of $\{\Delta, \mathsf{E} + \frac{n}{2}, \mathsf{F}\}$ that takes the standard basis $(X, H, Y) = \left(\left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right), \left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right), \left(\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right)\right)$ of $\mathfrak{sl}_2$ to $(\frac{1}{4\pi}\Delta, -(\mathsf{E}+\frac{n}{2}), -\pi\mathsf{F})$. The decomposition $\mathcal{P}_d = \bigoplus_{k=0}^{\lfloor d/2\rfloor} \mathcal{P}_d^k$ in Proposition 16ii then says in effect that $\mathcal{P} = \oplus_{d=0}^{\infty} \mathcal{P}_d^0 \otimes V_{\frac{n}{2}+d}$, where for any real $m > 0$ we write V_m for the infinite-dimensional irreducible representation of $\mathfrak{sl}_2$ with basis $\{Y^k v\}_{k=0}^{\infty}$ where $Xv = 0$ and $Hv = -mv$.

[19] The reader anxious to reach the application of weighted theta series to the study of extremal lattices etc. will likely want to skim or skip these remarks, and the last exercise for this section, at least on first reading(s).

Moreover, we noted already that $G_t = \exp(-\pi t \mathsf{F})$, and it is known that the Fourier transform is $e^{-\pi i n/4} \exp \frac{\pi i}{2}(\pi \mathsf{F} - \frac{1}{4\pi}\Delta)$.[20] Thus Lemmas 17 and 18, which give the action on $\Delta, \mathsf{E}, \mathsf{F}$ of conjugation by G_t and the Fourier transform, correspond to the action on $\mathfrak{sl}_2$ of conjugation by the elements $e^{tY} = \left(\begin{smallmatrix} 1 & 0 \\ t & 1 \end{smallmatrix}\right)$ and $e^{-\pi i(X+Y)/2} = -\left(\begin{smallmatrix} 0 & i \\ i & 0 \end{smallmatrix}\right)$ of SL_2.

In Proposition 19, we obtain $\hat{P}$ by applying to P the operators G_t, then the Fourier transform, then $G_{-1/t}$; up to the constant factor $e^{-\pi i n/4}$, this corresponds to the product

$$\begin{pmatrix} 1 & 0 \\ -1/t & 1 \end{pmatrix} \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} = -i \begin{pmatrix} t & 1 \\ 0 & -1/t \end{pmatrix},$$

in SL_2, which can be written as

$$-i \begin{pmatrix} t & 1 \\ 0 & -1/t \end{pmatrix} \begin{pmatrix} 1 & 1/t \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} t/i & 0 \\ 0 & (it)^{-1} \end{pmatrix} \begin{pmatrix} 1 & i/t \\ 0 & 1 \end{pmatrix}.$$

Now if $P \in \mathcal{P}_d^0$ then P is fixed by the one-parameter subgroup $\left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right)$ of SL_2 generated by X. Thus $\hat{P}$ is a multiple of the image of P under a diagonal matrix in SL_2. Such diagonal matrices constitute the one-parameter group generated by H. Since P is an eigenvector of H, we deduce that $\hat{P}$ is proportional to P. More precisely, since $HP = -(\frac{n}{2} + d)P$ we have $\mathrm{diag}(e^{\beta}, e^{-\beta})P = e^{\beta H}P = e^{-\beta(\frac{n}{2}+d)}P$; taking $e^{\beta} = t/i$ (with $\mathrm{Im}(\beta) = -\pi/2$),[21] and restoring the factor $e^{-\pi i n/4}$, we recover

$$\hat{P} = e^{-\pi i n/4}(t/i)^{-(\frac{n}{2}+d)}P = i^d t^{-(\frac{n}{2}+d)}P,$$

which is Theorem 12.

The differential operators $\partial^2/\partial x_j \partial x_k$, $x_j \partial/\partial x_k + \frac{1}{2}\delta_{jk}$, and $x_j x_k$ generate a Lie algebra isomorphic with $\mathfrak{sp}_{2n}$, which contains the span of $\Delta, \mathsf{E} + \frac{n}{2}, \mathsf{F}$ as the subalgebra invariant under the orthogonal group of $\mathbf{R}^n$. See the final exercise for this

[20] This has a memorable physical interpretation: running the Schrödinger equation on a quantum harmonic oscillator for $1/4$ of its classical period applies a multiple of the Fourier transform to the wave function. The distribution of factors of π in this formula is the reason we chose the homomorphism from $\mathfrak{sl}_2$ that maps $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ to $\frac{1}{4\pi}\Delta$ and $-\pi\mathsf{F}$, rather than $\frac{1}{2}\Delta$ and $-\frac{1}{2}\mathsf{F}$ which seems more natural at first.

Alternatively, we could use Lemma 18 to identify the Fourier transform with conjugation by $\pm i\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$.

[21] We must specify the path because in general a representation of $\mathfrak{sl}_2$ lifts to a representation not of SL_2 but of its universal cover. A more honest treatment would either carefully check that we are working in a contractible patch of $\mathrm{SL}_2(\mathbf{C})$, or compute the constant factor i^d of (81) in some other way.

section. Multiplying $\partial^2/\partial x_j \partial x_k$ and $x_j x_k$ by i (to make all the generators skew-Hermitian) yields infinitesimal generators of the Lie algebra of Weil's projective representation [Wei] of $\mathrm{Sp}_{2n}(\mathbf{R})$ on $L^2(\mathbf{R}^n)$.

Exercises

6.0 Verify directly that Theorem 21 holds for $n = 1$.

6.1 Check the commutation relations $[\mathsf{E}, \Delta] = -2\Delta$ and $[\mathsf{E}, \mathsf{F}] = +2\mathsf{F}$ of (90).

6.2 Verify the second formula in (96).

6.3 (Harmonic polynomials on $\mathbf{R}^2$) Let $n = 2$, and identify $\mathbf{R}^2$ with $\mathbf{C}$ as usual by $x_1 + ix_2 = z$. Show that for each $d \geq 1$ the polynomials z^d and $\bar{z}^d$ constitute a basis for $\mathcal{P}_d^0$. [Hint: recall that

$$\Delta = \Big(\frac{\partial}{\partial x_1} + i\frac{\partial}{\partial x_2}\Big)\Big(\frac{\partial}{\partial x_1} - i\frac{\partial}{\partial x_2}\Big).]$$

Deduce that the case $n = 2$ of Theorem 21 is tantamount to Weyl's characterization [Wey] of asymptotic equidistribution in $\mathbf{R}/\mathbf{Z}$, or equivalently in the unit circle $\Sigma_1 \cong \{z \in \mathbf{C} : |z| = 1\}$: a sequence of finite nonempty subsets $C_m \subset \Sigma_1$ is asymptotically equidistributed in Σ_1 if and only if for each $d = 1, 2, 3, \ldots$ we have $(\#C_m)^{-1} \sum_{z \in C_m} z^d \to 0$ as $k \to \infty$.[22]

6.4 (a weighted theta function twisted by a Dirichlet character) Show that if f is a Schwartz function on $\mathbf{R}$ then

$$\sum_{m=-\infty}^{\infty} \chi_4(m) f(m) = \frac{1}{2i} \sum_{n=-\infty}^{\infty} \chi_4(n) \hat{f}(n/4),$$

where χ_4 is the Dirichlet character mod 4 introduced in Exercise 3.7. Since χ_4 is odd, we would get only the trivial identity $0 = 0$ using the even function $f(x) = e^{-\pi t x^2}$. However $f(x) = xe^{-\pi t x^2}$ yields a new identity. Use this identity to prove that

$$\sum_{n=1}^{\infty} n\chi_4(n) q^{n^2/8} = q^{1/8}(1 - 3q + 5q^3 - 7q^6 + 9q^{10} - 11q^{15} + - \cdots) \quad (106)$$

is a modular form of weight $3/2$ for Γ and the 8th roots of unity $\varepsilon_{c,d}^3$, with $\varepsilon_{c,d}$ as in Exercise 2.5; deduce that this modular form is η^3.

[22] Weyl also shows that that this is equivalent to the condition that $\#(I \cap C_m)/\#C_m \to |I|/|\Sigma|$ for all arcs $I \subseteq \Sigma$ (where $|\cdot|$ denotes the length). See also [Kö, Ch.3]. Analogous statements can be made on the unit sphere in $\mathbf{R}^n$ for any n, but we do not need those versions of equidistribution here.

6.5 Prove that there are positive constants $c_0, c_2, c_4, \ldots$ (depending on n) such that if φ is the restriction to Σ of a homogeneous polynomial P of even degree d then

$$\int_{x\in\Sigma} \varphi(x)\, d\nu_x = c_d t^{(n+d)/2} \int_{x\in\mathbf{R}^n} e^{-\pi t\langle x,x\rangle}\, P(x)\, d\mu_x. \tag{107}$$

for all $t > 0$. We defined the integral over Σ so that $c_0 = 1$; can you evaluate the other c_d explicitly? (For odd d, (107) holds for all c_d because both sides vanish...)

6.6 Define an inner product $\langle \cdot, \cdot \rangle$ on the space of Schwartz functions on $\mathbf{R}^n$ by

$$\langle f, g\rangle := \int_{x\in\mathbf{R}^n} f(x)\, \overline{g(x)}\, d\mu_x. \tag{108}$$

Recall that an operator A on this space is *self-adjoint* if $\langle Af, g\rangle = \langle f, Ag\rangle$ for all Schwartz functions f, g. Prove that the operators Δ and F are self-adjoint. [This is immediate for F, and uses integration by parts for Δ. Warning: E is *not* self-adjoint; indeed the fact that Δ and F are self-adjoint implies that $4\mathsf{E} + 2n = [\Delta, \mathsf{F}]$ is anti-self-adjoint.] Use this and the Corollary to Lemma 17 to prove that $\mathcal{P}^0_d$ and $\mathcal{P}^0_{d'}$ are orthogonal for $d \neq d'$, that is, that $\int_{x\in\Sigma} P(x)\, Q(x)\, d\nu_x = 0$ if P and Q are harmonic polynomials of different degrees. (We showed the special case $d = 0$ during the proof of Theorem 21.) Recover the orthogonal decomposition $L^2(\Sigma, d\nu_x) = \bigoplus_{d=0}^{\infty} \mathcal{P}^0_d$ using this and Proposition 20.

6.7 (Failure of equidistribution for $n = 2$) Equidistribution of $S_{k_m}(L)$ can fail for $n = 2$ even if $N_{k_m}(L) \to \infty$, and already for the familiar lattice $L = \mathbf{Z}^2$. To study this question we must use the intimate connection between the Diophantine equation $\langle x, x\rangle = x_1^2 + x_2^2 = k$ and arithmetic in the Gaussian integers $\{x_1 + ix_2 \mid x_1, x_2 \in \mathbf{Z}\}$ that was kept in the background of Exercises 3.7 and 3.8.

(i) Let p be a prime such that $p \equiv 1 \bmod 4$, and $\{k_m\}$ a sequence of integers with $p^m | k_m$. Then if $N_{k_m} \neq 0$ then $N_{k_m} > 4m$; in particular $N_{k_m} \to \infty$. Show that in this case $\{S_{k_m}\}$ is asymptotically equidistributed on the unit circle.

(ii) Prove that there exist k_m such that $N_{k_m} \to \infty$ but $\{S_{k_m}\}$ is *not* asymptotically equidistributed on the unit circle. [Use the following special case of Hecke's theorem [H1] on equidistribution of prime ideals in number fields: for all ϵ there exist integers x_1, x_2 with $0 < x_2 < \epsilon x_1$ such that $x_1^2 + x_2^2$ is prime.]

6.8 (Identification of a Lie algebra of differential operators with $\mathfrak{sp}_{2n}$) Let V be the $2n$-dimensional $\mathbf{R}$-vector space of differential operators on $\mathbf{R}^n$ with basis x_j and $\partial/\partial x_j$ ($1 \leq j \leq n$); and let $\mathfrak{g}$ be the space of dimension $2n^2 + n$ with basis $\partial^2/\partial x_j \partial x_k$, $x_j \partial/\partial x_k + \frac{1}{2}\delta_{jk}$, and $x_j x_k$ ($1 \leq j, k \leq n$).
i) Check that if $A, B \in \mathfrak{g}$ then $[A, B] \in \mathfrak{g}$, while if $A \in \mathfrak{g}$ and $v \in V$ then $[A, v] \in V$. Thus $\mathfrak{g}$ is a Lie algebra and V is a representation of $\mathfrak{g}$.

ii) The map $[\cdot,\cdot] : V \times V \to \mathbf{R}$ is a perfect alternating pairing on V. Use the Jacobi identity $[[A,B],C] + [[B,C],A] + [[C,A],B] = 0$ to prove that $[gv, v'] + [v, gv'] = 0$ for all $g \in \mathfrak{g}$ and $v, v' \in V$. Thus $\mathfrak{g}$ is contained in the Lie algebra of the symplectic group for our pairing; since $\dim \mathfrak{g} = \dim \mathfrak{sp}_{2n}$, this gives an isomomorphism $\mathfrak{g} \xrightarrow{\sim} \mathfrak{sp}_{2n}$.

7. Extremal lattices and spherical designs

8. Further directions

Lattices with arbitrary discriminants; periodic weights

Higher θ functions; coding analogues

Murphy's Law

Acknowledgements

References

[Bl] Blcihfeldt, H.F.: The minimum values of positive quadratic forms in six, seven and eight variables. *Math. Z.* **39** (1935), 1–15.

[Bo] Borcherds, R.E.: Table -2: the norm -2 vectors in $\mathrm{II}_{25,1}$. Online at http://math.berkeley.edu/∼reb/lattices/table2.html .

[Coh] Cohen, H.: *Number theory, Vol. II: Analytic and Modern Tools*, New York: Springer, 2007 (GTM **240**).

[CE] Cohn, H., and Elkies, N.D.: New upper bounds on sphere packings I, *Annals of Math.* **157** (2003), 689–714 arXiv:math/0110009 [math.MG]).

[CK] Cohn, H., and Kumar, A.: Optimality and uniqueness of the Leech lattice among lattices, to appear in the *Annals of Math.* (arXiv:math/0403263 [math.MG]).

[Con] Conway, J. H.: A characterisation of Leech's lattice, *Invent. Math.* **7** (1969), 137–142.

[CS1] Conway, J.H., and Sloane, N.J.A.: A new upper bound for the minimum of an integral lattice of determinant one, *Bull. Amer. Math. Soc.* **23** (1990), 383–387; Erratum: **24** (1991), 479.

[CS2] Conway, J.H., and Sloane, N.J.A.: *Sphere Packings, Lattices and Groups* (3rd ed.). New York: Springer 1999.

[Di] Dickson, L.E.: *History of the Theory of Numbers, Vol. II: Diophantine Analysis.* New York: Stechert & Co., 1934.

[Ed] Edwards, H.M.: *Riemann's zeta function.* New York: Academic Press, 1974.

[E1] Elkies, N.D.: A characterization of the $\mathbf{Z}^n$ lattice, *Math. Research Letters* **2** (1995), 321–326 (arXiv:math/9906019v1 [math.NT]).

[E2] Elkies, N.D.: Lattices and codes with long shadows, *Math. Research Letters* **2** (1995), 643–651 (arXiv:math/9906086v1 [math.NT]).

[E3] Elkies, N.D.: Lattices, Linear Codes, and Invariants, *Notices of the American Math. Soc.* **47** (2000), 1238–1245 and 1382–1391.

[EV] Ellenberg, J., and Venkatesh, A.: Local-global principles for representations of quadratic forms, *Invent. Math.* **171** #2 (2008), 257–279.

[H1] Hecke, E.: Eine neue Art von Zetafunktionen und ihre Beziehungen zur Verteilung der Primzahlen II, *Math. Z.* **6** (1920), 11–51.

[H2] Hecke, E.: Analytische Arithmetik der positiven quadratischen Formen, *Kgl. Danske Vid. Selsk. Math.-Fys. Medd.* **17** (1940) #12 [= pages 789–918 in *Mathematische Werke*, Göttingen: Vandenhoeck & Ruprecht, 1959].

[Iw] Iwaniec, H.: *Topics in Classical Automorphic Forms*. Providence, RI: American Math. Society, 1997.

[KD] Kaba, M., "in collaboration with" Dickson, L.E.: On the Representation of Numbers as the Sum of Two Squares, *American Math. Monthly* **16** #5 (May 1909), 85–87.

[Ka] Kac, M.: Can one hear the shape of a drum? *American Math. Monthly* **73** (1966), 1–23.

[KAL] Klosinski, L.F., Alexanderson, G.L., and Larson, L.C.: The Sixty-Sixth William Lowell Putnam Mathematical Competition, *American Math. Monthly* **113** #8 (October 2006), 733–743.

[Ki] King, O.D.: A mass formula for unimodular lattices with no roots, *Math. Comp.* **72** (2003), 839–863 (arXiv:math/0012231v1 [math.NT]). Online tables at http://math.berkeley.edu/∼reb/lattices/table.txt .

[Kö] Körner, T.W.: *Fourier Analysis.* Cambridge, England: Cambridge University Press, 1988.

[Le] Leech, J.: `construction of Leech lattice, c. 1960`

[Mu] Munroe, R.: E TO THE PI MINUS PI, *xkcd* **217** (2007) (http://www.xkcd.com/217/).

[MOS] Mallows, C.L., Odlyzko, A.M., and Sloane, N.J.A.: Upper Bounds for Modular Forms, Lattices, and Codes. *J. Alg.* **36**, 68–766 (1975).

[Min] Minkowski, H.: Grundlagen für eine Theorie der quadratischen Formen mit ganzzahligen Koeffizienten, pages 3–145 in *Gesammelte Abhandlungen* (Leipzig, 1991, republished 1967 by Chelsea, New York); German translation of "Mémoire sur la théorie des formes quadratiques à coefficients entiers", *Mémoirs présentés par divers savants à l'Academie des Sciences de l'Institut de France* **29** (1887), 1–180.

[Mil] Milnor, J.: Eigenvalues of the Laplace operator on certain manifolds, *Proc. Nat. Acad. Sci. USA* **51** (1964), 542.

[Ni] Niemeier, H.-V.: Definite quadratische Formen der Dimension 24 und Diskriminante 1, em J. Number Theory **5** (142–178), 1973.

[Po] Pommerenke, C: Über die Gleichverteilung von Gitterpunkten auf m-dimensionalen Ellipsoiden, *Acta Arith.* **5** (1959), 227–257.

[RS] Rains, E.M., and Sloane, N.J.A.: The Shadow Theory of Modular and Unimodular Lattices, *J. Number Theory* **73** (1998), 359–389 [= http://www.research.att.com/$\sim$njas/doc/mod.pdf].

[SP] Schulze-Pillot, R.: Representation by integral quadratic forms — a survey. Pages 303–321 in *Algebraic and arithmetic theory of quadratic forms* (Contemp. Math.) **344**, Providence, RI: Amer. Math. Soc., 2004.

[Sch] Schoeneberg, B.: Das Verhalten von mehrfachen Thetareihen bei Modulsubstitutionen, *Math. Annalen* **116** (1939), 511–523.

[Se] Serre, J.-P.: *A Course in Arithmetic.* New York: Springer, 1973.

[Si] Siegel, C.L.: Berechnung von Zetafunktionen an ganzzahligen Stellen. *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II* **1969**, 87–102 (1969) [= pages 82–97 in *Gesammelte Abhandlungen IV*, Berlin: Springer 1979].

[SS] Stein, E.M., and Shakarchi, R.: *Complex Analysis.* Princeton, NJ: Princeton University Press, 2003.

[SW] Stein, E.M., and Weiss, G.L.: *Introduction to Fourier Analysis on Euclidean Spaces, Princeton Math. Series* **32**, Princeton, NJ: Princeton University Press, 1971.

[Ve] Venkov, B.B.: On the classification of integral even unimodular 24-dimensional quadratic forms. *Trudy Mat. Inst. Steklov.* **148** 65–76 (1978). [In Russian; trans. by the Amer. Math. Soc. as *Proc. Steklov Inst. Math.* **148** 63–74 (1980); also the source of [CS2, Ch.18]]

[Wei] Weil, A.: Sur certaines groupes d'opérateurs unitaires. *Acta Math.* **111**, 143–211 (1964).

[Wey] Weyl, H.: Über ein Problem aus dem Gebiete der diophantischen Approximationen, *Nachrichten der Königlichen Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse*, 1914, 234–244 [= *Gesammelte Abhandlungen I* (Springer: Berlin 1968), 487–497].